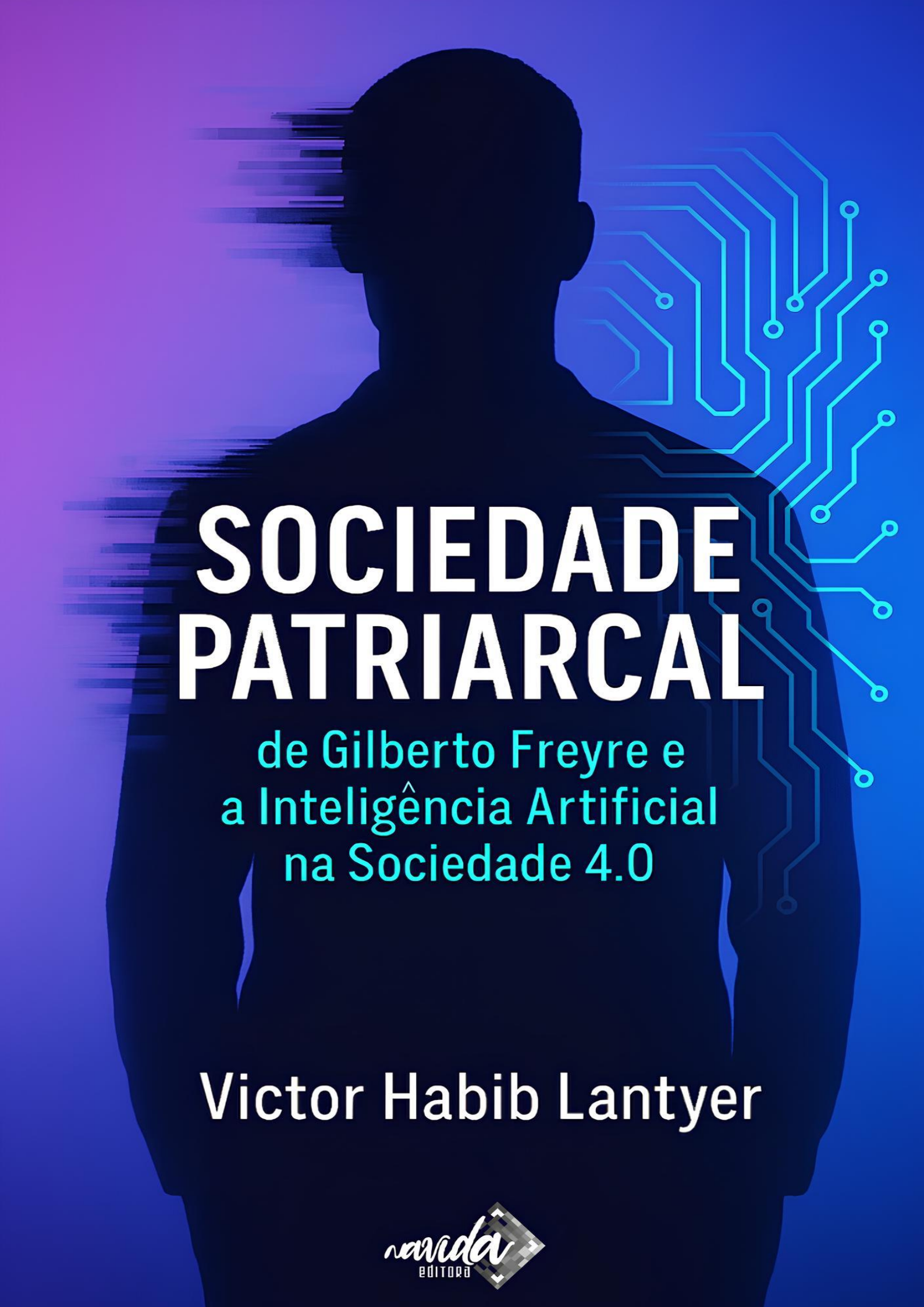
SOCIEDADE PATRIARCAL
de Gilberto Freyre e a Inteligência Artificial na Sociedade 4.0
Victor Habib Lantyer
navida editora

Copyright© 2025 Navida Editora

Nenhuma parte deste livro pode ser reproduzida, sejam quais forem os meios empregados, a não ser com a permissão escrita do autor e da editora, conforme a Lei nº 9610, de 19 de fevereiro de 1998.

Título: A Sociedade Patriarcal de Gilberto Freyre e a Inteligência Artificial: Uma Reflexão Crítica

Autor: Victor Habib Lantyer

Coordenação editorial: Laíze Lantyer Luz

Capa: Victor Habib Lantyer

Diagramação: Victor Habib Lantyer

Revisão: Laíze Lantyer Luz

Conselho editorial:

Doutores:

Aída Cristina do Nascimento Silva

André Alves Portella

Antonio Pedro Silva

Bárbara Pontes

Dalzimar Andrade

Elias Sampaio

Isabel Colucci Coelho

Laura Braz

Milton Vasconcellos

Ricardo Xavier

Silvana Sá de Carvalho

Taís Dórea

Wdileston Batista

Mestres:

Alexandre Sion

Ana Paula Mendes Geitenes

Diogo Assis Cardoso Guanabara

Isis Veiga

Joilson Santana

Lisdeili Nobre

Paulo Cezar Martins Pinto

Sara Meireles

www.navidaeditora.com
contato@navidaeditora.com
@navidaeditora
doi.org/10.51968/navida.978-65-89020-22-6

Dados Internacionais de Catalogação na Publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)

Lantyer, Victor Habib
A sociedade patriarcal de Gilberto Freyre e a inteligência artificial na sociedade [livro eletrônico] : uma reflexão crítica / Victor Habib Lantyer. -- 1. ed. -- Salvador, BA : NAVIDA, 2025.
ePDF

Bibliografia
ISBN 978-65-89020-22-6

1. Inteligência artificial 2. Freyre, Gilberto, 1900-1987 - Crítica e interpretação 3. Patriarcado I. Título.

25-288911 CDD-981

Índices para catálogo sistemático:

1. Freyre, Gilberto : Ensaios : História 981

Maria Alice Ferreira - Bibliotecária - CRB-8/7964

PREFÁCIO

A interrogação que deu origem a este ensaio surgiu de uma inquietação, de um aparente anacronismo que, ao ser investigado, revelou-se uma perturbadora continuidade. Como poderiam as estruturas arcaicas de uma sociedade patriarcal e escravocrata, dissecadas com maestria por Gilberto Freyre há quase um século, encontrar um eco, uma ressonância, na arquitetura asséptica e ultracontemporânea da inteligência artificial? A tarefa a que me propus nestas páginas foi a de demonstrar que tal conexão não é nem fortuita nem superficial.

Ao mergulhar novamente em Casa-Grande & Senzala, percebi que a análise freyriana transcendia a descrição histórica para oferecer um modelo de poder — capilar, doméstico, mas fundamentalmente estruturante. O patriarcado, em Freyre, é menos uma relíquia do passado e mais uma gramática das relações sociais brasileiras. Foi essa gramática que busquei identificar nos códigos e algoritmos que hoje governam desde a seleção de um currículo até a concessão de crédito.

Este livro não é, portanto, um mero exercício de justaposição teórica. É uma tentativa de forjar uma lente crítica para o presente, utilizando o instrumental analítico de Freyre, enriquecido por aportes do materialismo histórico, da crítica da razão instrumental da Escola de Frankfurt e do pensamento feminista, para iluminar as novas formas de dominação que se erguem sob o signo da inovação tecnológica.

A jornada que proponho ao leitor parte da desconstrução dos conceitos fundamentais de Freyre e da IA para, em seguida, evidenciar como os vieses algorítmicos se tornam os novos perpetuadores de velhas hierarquias. Contudo, não me detenho na denúncia. Inspirado na própria noção freyriana de "plasticidade cultural", dedico a parte final a explorar as fissuras neste sistema, a investigar como a mesma tecnologia que reproduz a opressão pode, paradoxalmente, ser mobilizada como ferramenta de transformação e equidade.

Se a casa-grande foi o epicentro de uma ordem social, os oligopólios tecnológicos e seus algoritmos opacos configuram um novo centro de poder. Compreender suas dinâmicas, seus pressupostos e seus impactos é, a meu ver, uma das tarefas intelectuais mais urgentes do nosso tempo. Este ensaio é o meu modesto contributo para essa empreitada.

VICTOR HABIB LANTYER

SUMÁRIO

1. Introdução 1
2. Conceitos Fundamentais de Gilberto Freyre 3
3. Fundamentos da Inteligência Artificial 10
4. Viés e Discriminação Algorítmica 21
5. A Reprodução de Estruturas Patriarcais pela IA 24
6. Patriarcalismo e Oligopólios Tecnológicos: A Concentração de Poder na Era dos Dados 30
7. Casos de Reprodução de Vieses pela IA 34
   7.1. O Algoritmo de Recrutamento da Amazon 34
   7.2. O Algoritmo COMPAS no Sistema Judicial 36
   7.3. Sistemas de Reconhecimento Facial e Vieses Raciais e de Gênero 38
   7.4. Google Translate e o Reforço de Estereótipos 40
   7.5. Algoritmos em Plataformas de Mídia Social 41
   7.6. Sistemas de Avaliação de Crédito e Discriminação Financeira 43
   7.7. Chatbots e Discurso Discriminatório 46
8. Abordagens Críticas: Materialismo Histórico-Dialético, Escola de Frankfurt, Feminismo e Sociologia da Ciência e Tecnologia 47
   8.1. Materialismo Histórico-Dialético 48
   8.2. A Escola de Frankfurt e a Crítica da Razão Instrumental 51
   8.3. Feminismo e Teoria de Gênero: Desconstruindo Vieses e Discriminações na IA 56
   8.4. Sociologia da Ciência e Tecnologia 57
   8.5. Construção Social da Tecnologia e a Inteligência Artificial 60
9. Quebrando a Estrutura: IA como Ferramenta de Transformação 63
   9.1. A Importância da Plasticidade Cultural 64
   9.2. Convivência de Contrastes e Harmonia nas Diferenças 65
   9.3. A Família Patriarcal e a Modernização Tecnológica 66

9.4. Regionalismo e Valorização das Culturas Locais 68
9.5. Educação e Inteligência Artificial como Ferramentas de Transformação Social 72
10. Estratégias Práticas para Melhorar Dados de Treinamento de IA e Reduzir Vieses Discriminatórios 78
10.1. Representatividade nos Dados de Treinamento 78
10.2. Processos de Anotação de Dados e Vieses Humanos 80
10.3. Identificação e Correção de Vieses nos Dados 81
10.4. Desenvolvimento de Modelos com Consideração de Equidade 83
10.5. Transparência e Auditoria nos Processos de IA 84
10.6. Educação e Sensibilização sobre Vieses em IA 85
10.7. Políticas e Regulamentações para Práticas Responsáveis 86
11. Conclusão 87
Referências 91

1. Introdução

Este ensaio teórico tem como objetivo analisar como as estruturas patriarcais descritas por Gilberto Freyre que podem estar sendo reproduzidas e perpetuadas pela inteligência artificial (IA) na sociedade contemporânea.[1] Freyre traçou um retrato profundo das relações de poder no Brasil colonial e pós-colonial, destacando o papel central do patriarcalismo na formação social e familiar.[2] O patriarca, figura masculina dominante, detinha controle sobre a economia, a família e as decisões políticas, perpetuando uma estrutura hierárquica onde a mulher, o escravizado e os filhos ocupavam papéis de submissão.[3] Essa ordem patriarcal não se limitava apenas à esfera doméstica, mas também servia como microcosmo das relações de poder na sociedade em geral, refletindo uma dominação masculina estrutural e sistêmica.[4]

O conceito de patriarcado em Freyre não se restringe ao espaço privado, mas permeia as estruturas de poder público e econômico.[5] Ele estabelece uma continuidade histórica entre o Brasil colonial e o Brasil contemporâneo, evidenciando que as hierarquias patriarcais ainda sustentam as relações de poder na sociedade atual.[6] Nesse contexto, a modernidade não eliminou essas estruturas; ao contrário, elas se adaptaram e se manifestam em novas formas, incluindo as tecnologias emergentes.[7]

Com o advento da inteligência artificial, uma tecnologia que está rapidamente se integrando aos processos de tomada de decisão em diversas

[1] FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).
[2] FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).
[3] HOLANDA, S. B. de. Raízes do Brasil. São Paulo: Companhia das Letras, 1995.
[4] FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).
[5] FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).
[6] ORTIZ, R. Cultura brasileira e identidade nacional. São Paulo: Brasiliense, 2002.
[7] ORTIZ, R. Cultura brasileira e identidade nacional. São Paulo: Brasiliense, 2002.

áreas, surgem novos desafios para essas dinâmicas históricas.[8] A IA não é neutra; ela é desenvolvida, alimentada e treinada por seres humanos que operam dentro de contextos culturais, históricos e sociais, muitas vezes refletindo os vieses e estruturas de poder preexistentes.[9] Isso levanta a questão central deste ensaio: a IA está reproduzindo o patriarcalismo que Freyre descreveu?

Além disso, é crucial considerar como a falta de diversidade nas equipes de desenvolvimento de IA influencia os resultados produzidos por essas tecnologias. A predominância masculina nas áreas de ciência da computação e engenharia contribui para a manutenção de perspectivas patriarcais na concepção e implementação de sistemas de IA.[10] Essa homogeneidade pode levar à criação de algoritmos e modelos que não consideram adequadamente as necessidades e experiências de grupos marginalizados, incluindo mulheres e minorias étnicas.[11]

Portanto, este ensaio busca não apenas identificar a presença de estruturas patriarcais na IA, mas também discutir possíveis caminhos para mitigar a reprodução desses vieses. Em última instância, o objetivo é contribuir para o debate sobre como a IA pode ser desenvolvida e implementada de maneira a promover uma sociedade mais justa e equitativa, alinhada com os avanços nas discussões sobre igualdade de gênero e justiça social.

---

[8] RUSSELL, S.; NORVIG, P. Artificial Intelligence: A Modern Approach. 3. ed. Upper Saddle River: Prentice Hall, 2016.
[9] NOBLE, S. U. Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press, 2018.
[10] HICKS, M. Programmed Inequality: How Britain Discarded Women Technologists and Lost Its Edge in Computing. Cambridge: MIT Press, 2017.
[11] CRAWFORD, K.; CALO, R. There is a blind spot in AI research. Nature, v. 538, n. 7625, p. 311-313, 2016.

2. Conceitos Fundamentais de Gilberto Freyre

Para compreender a reprodução de preconceitos estruturais, como as estruturas patriarcais, raciais e de classe na inteligência artificial, é essencial analisar os principais conceitos desenvolvidos por Gilberto Freyre em Casa-Grande & Senzala e em suas obras subsequentes.[12] Freyre ofereceu uma interpretação inovadora da formação da sociedade brasileira, destacando a complexidade das relações sociais, raciais e culturais que emergiram durante o período colonial.[13]

Gilberto Freyre nasceu em 15 de março de 1900, em Recife, Pernambuco.[14] Educado em ciências sociais nos Estados Unidos, foi influenciado pelo movimento culturalista da antropologia norte-americana, especialmente pelas ideias de Franz Boas.[15] Ao retornar ao Brasil, dedicou-se ao estudo da sociedade brasileira, buscando compreender as especificidades da formação nacional.[16] Sua obra mais conhecida, Casa-Grande & Senzala, publicada em 1933, resultou de uma profunda pesquisa histórica e antropológica sobre o período colonial.[17]

A escrita de Sobrados e Mucambos, publicada em 1936, foi motivada pelo desejo de Freyre de dar continuidade à análise iniciada em Casa-Grande & Senzala (Freyre, 1936).[18] Enquanto esta última focalizava a formação rural e patriarcal da sociedade brasileira, Sobrados e Mucambos buscava compreender as transformações ocorridas com a urbanização e a transição para o período

[12] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933
[13] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933
[14] FREYRE, G. Tempo morto e outros tempos. Rio de Janeiro: José Olympio, 1981.
[15] FREYRE, G. Tempo morto e outros tempos. Rio de Janeiro: José Olympio, 1981.
[16] FREYRE, G. Tempo morto e outros tempos. Rio de Janeiro: José Olympio, 1981.
[17] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933
[18] FREYRE, G. Sobrados e Mucambos. Rio de Janeiro: José Olympio, 1936.

pós-colonial.[19] O contexto histórico do início do século XX, marcado por processos de modernização, industrialização e migração interna, influenciou Freyre a explorar como essas mudanças afetaram as estruturas sociais herdadas do passado colonial.[20]

No período em que Freyre escreveu Sobrados e Mucambos, o Brasil passava por significativas transformações políticas e sociais.[21] A Revolução de 1930 havia derrubado a Primeira República, levando Getúlio Vargas ao poder e inaugurando uma era de reformas sociais e econômicas.[22] Esse contexto de mudança inspirou Freyre a investigar como as estruturas patriarcais e hierárquicas se adaptaram ao novo cenário urbano e moderno.[23] Ele buscou compreender se as transformações em curso estavam realmente alterando as relações de poder ou apenas reconfigurando-as em novas formas.[24]

Em Sobrados e Mucambos, Freyre analisa a transição da sociedade rural para a urbana, destacando como os sobrados—residências urbanas da elite—substituíram as casas-grandes rurais, enquanto os mucambos—habitações precárias das classes populares—sucederam as senzalas.[25] Ele observa que, apesar das mudanças na paisagem física e nas formas de habitação, as hierarquias sociais e a concentração de poder nas mãos de poucos persistiram.[26] Essa obra aprofunda a compreensão das continuidades e rupturas na formação

[19] FREYRE, G. Sobrados e Mucambos. Rio de Janeiro: José Olympio, 1936.
[20] NOGUEIRA, O. Preconceito de marca: as relações raciais em Itapetininga. São Paulo: Edusp, 2003.
[21] FREYRE, G. Sobrados e Mucambos. Rio de Janeiro: José Olympio, 1936.
[22] FAUSTO, B. História do Brasil. São Paulo: Edusp, 1995.
[23] FREYRE, G. Sobrados e Mucambos. Rio de Janeiro: José Olympio, 1936.
[24] NOGUEIRA, O. Preconceito de marca: as relações raciais em Itapetininga. São Paulo: Edusp, 2003.
[25] FREYRE, G. Sobrados e Mucambos. Rio de Janeiro: José Olympio, 1936.
[26] FREYRE, G. Sobrados e Mucambos. Rio de Janeiro: José Olympio, 1936.

social brasileira, evidenciando a resiliência das estruturas patriarcais mesmo diante da modernização.[27]

Em Casa-Grande & Senzala, Freyre introduz o conceito de patriarcado rural como a base da organização social brasileira.[28] O patriarca, senhor de engenho, exercia autoridade absoluta não apenas sobre sua família, mas também sobre os escravizados e agregados, estabelecendo uma hierarquia rígida e centralizada.[29] Essa estrutura patriarcal era sustentada pela economia açucareira e pela grande propriedade rural, a casa-grande, que simbolizava o poder e a opulência do senhor.[30]

Esse patriarcado rural não se limitava à gestão econômica, mas permeava todos os aspectos da vida social e cultural.[31] O senhor de engenho era responsável por administrar a justiça local, atuar como intermediário com autoridades coloniais e influenciar práticas religiosas.[32] A educação dos filhos e a moralidade da família eram diretamente controladas pelo patriarca, reforçando a transmissão intergeracional das estruturas de poder.[33] Assim, o patriarcado rural estabeleceu as bases para uma cultura de autoritarismo e centralização que perdura em várias formas na sociedade brasileira contemporânea.[34]

A senzala, por sua vez, representava o espaço dos escravizados, um local de submissão e resistência cultural.[35] Freyre argumenta que, apesar da opressão, os escravizados africanos e seus descendentes contribuíram

[27] FREYRE, G. Sobrados e Mucambos. Rio de Janeiro: José Olympio, 1936.
[28] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[29] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[30] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[31] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[32] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[33] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[34] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[35] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.

significativamente para a formação cultural brasileira, através de um processo de miscigenação e sincretismo.[36] Essa interação entre a casa-grande e a senzala foi fundamental para a construção da identidade nacional, marcada pela diversidade racial e cultural.[37]

A relação entre a casa-grande e a senzala não era apenas física, mas também simbólica.[38] Enquanto a casa-grande representava o poder e a dominação, a senzala era o espaço da resistência e da preservação de práticas culturais africanas.[39] As manifestações culturais, religiosas e linguísticas dos escravizados influenciaram profundamente a cultura brasileira, apesar das tentativas de supressão por parte dos colonizadores.[40] Esse processo resultou em um sincretismo que é uma característica distintiva da identidade nacional.[41]

Freyre também destaca o papel ambíguo da mulher na sociedade patriarcal.[42] Enquanto as mulheres brancas da elite eram confinadas ao espaço doméstico e submetidas à autoridade masculina, as mulheres negras e indígenas eram frequentemente exploradas sexualmente, resultando em relações complexas de poder.[43] Essa dinâmica contribuiu para a formação de uma sociedade mestiça, mas também perpetuou a subordinação feminina em múltiplos níveis.[44]

As mulheres desempenhavam papéis fundamentais na manutenção da estrutura patriarcal, embora frequentemente invisibilizadas nas narrativas

[36] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[37] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[38] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[39] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[40] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[41] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[42] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[43] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[44] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.

históricas.[45] As mulheres brancas eram educadas para serem esposas submissas e mães dedicadas, reforçando a continuidade do poder masculino.[46] Por outro lado, as mulheres negras e indígenas sofriam duplas opressões: de gênero e de raça.[47] A exploração sexual dessas mulheres não apenas servia aos desejos dos senhores, mas também era um mecanismo de dominação e controle social.[48]

A ideia de "democracia racial" é outro conceito-chave em Freyre.[49] Ele sugere que, apesar das tensões raciais, o Brasil desenvolveu relações mais harmoniosas entre diferentes grupos étnicos em comparação com outros países.[50] No entanto, críticos argumentam que essa visão romantiza as desigualdades e invisibiliza as injustiças sofridas por negros e indígenas.[51]

A crítica ao mito da democracia racial destaca que a suposta harmonia esconde profundas desigualdades e discriminações estruturais.[52] Estudos sociológicos e antropológicos demonstram que o racismo no Brasil é sutil e permeia diversas esferas da sociedade, desde o mercado de trabalho até o sistema educacional.[53] Assim, a ideia de democracia racial pode servir para desviar a atenção das necessidades urgentes de políticas antirracistas efetivas.[54]

Freyre também aborda a adaptação cultural como um processo central na formação brasileira.[55] A sociedade colonial não replicou simplesmente os modelos europeus, mas os adaptou às condições locais, resultando em novas

---

[45] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[46] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[47] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[48] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[49] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[50] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[51] MUNANGA, K. Rediscutindo a Mestiçagem no Brasil: Identidade Nacional versus Identidade Negra. Petrópolis: Vozes, 1999.
[52] MUNANGA, K. Rediscutindo a Mestiçagem no Brasil: Identidade Nacional versus Identidade Negra. Petrópolis: Vozes, 1999.
[53] HASENBALG, C. Discriminação e Desigualdades Raciais no Brasil. Rio de Janeiro: Graal, 1979.
[54] MUNANGA, K. Rediscutindo a Mestiçagem no Brasil: Identidade Nacional versus Identidade Negra. Petrópolis: Vozes, 1999.
[55] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.

formas de organização social e cultural.[56] Essa perspectiva ressalta a criatividade e a resiliência das populações locais, mas também reconhece as imposições do poder colonial.[57]

A adaptação cultural envolveu não apenas a incorporação de elementos indígenas e africanos, mas também a transformação de práticas europeias para se adequar ao clima, à geografia e às circunstâncias sociais do Brasil.[58] Isso resultou em uma cultura híbrida, na qual elementos distintos se fundiram para criar algo novo e único.[59] No entanto, esse processo também foi marcado por assimetrias de poder, onde a cultura europeia muitas vezes era privilegiada em detrimento das demais.[60]

A relação entre modernidade e tradição é explorada em Ordem e Progresso, onde Freyre analisa as transformações políticas e sociais no Brasil republicano.[61] Ele argumenta que as influências estrangeiras e o processo de industrialização não eliminaram as estruturas patriarcais, mas as reconfiguraram em novos contextos.[62] Isso sugere uma continuidade histórica das relações de poder.[63]

A modernização brasileira frequentemente adotou modelos externos sem considerar as especificidades locais, levando a tensões entre progresso e identidade nacional.[64] Freyre aponta que, apesar das mudanças aparentes, as elites mantiveram o controle político e econômico, adaptando as estruturas

[56] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[57] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[58] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[59] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[60] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[61] FREYRE, G. Ordem e Progresso. Rio de Janeiro: José Olympio, 1959.
[62] FREYRE, G. Ordem e Progresso. Rio de Janeiro: José Olympio, 1959.
[63] FREYRE, G. Ordem e Progresso. Rio de Janeiro: José Olympio, 1959.
[64] BOSI, A. Dialética da Colonização. São Paulo: Companhia das Letras, 1992.

patriarcais para se adequar aos novos tempos.[65] Isso evidencia a resiliência dessas estruturas e a necessidade de uma análise crítica sobre como elas persistem.[66]

O conceito de "lusotropicalismo" é outra contribuição de Freyre, enfatizando a capacidade dos portugueses de se adaptarem aos trópicos e de se misturarem com outras culturas.[67] Embora tenha sido utilizado para justificar políticas coloniais, esse conceito também destaca a complexidade das identidades culturais no contexto brasileiro.[68]

O lusotropicalismo foi controverso, pois serviu tanto para celebrar a diversidade cultural quanto para legitimar a dominação portuguesa em suas colônias.[69] No contexto brasileiro, o conceito ajudou a explicar a formação de uma sociedade mestiça, mas também foi criticado por minimizar as violências do colonialismo.[70] Compreender o lusotropicalismo é fundamental para analisar como narrativas históricas podem ser utilizadas para sustentar estruturas de poder.[71]

A centralidade da família patriarcal na organização social é um tema recorrente.[72] Freyre mostra como a família extensa, liderada pelo patriarca, funcionava como unidade econômica, social e política.[73] Essa estrutura influenciou as relações sociais, as práticas culturais e até mesmo as instituições políticas, perpetuando uma cultura de autoritarismo e paternalismo.[74]

---

[65] FREYRE, G. Ordem e Progresso. Rio de Janeiro: José Olympio, 1959.
[66] FREYRE, G. Ordem e Progresso. Rio de Janeiro: José Olympio, 1959.
[67] FREYRE, G. O Mundo que o Português Criou. Lisboa: Livros do Brasil, 1952.
[68] FREYRE, G. O Mundo que o Português Criou. Lisboa: Livros do Brasil, 1952.
[69] FREYRE, G. O Mundo que o Português Criou. Lisboa: Livros do Brasil, 1952.
[70] FERNANDES, F. A integração do negro na sociedade de classes. São Paulo: Ática, 1978.
[71] FREYRE, G. O Mundo que o Português Criou. Lisboa: Livros do Brasil, 1952.
[72] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[73] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[74] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.

A família patriarcal serviu como microcosmo da sociedade, onde as relações de poder eram reproduzidas e naturalizadas.[75] O patriarca não apenas detinha poder sobre sua família, mas também desempenhava um papel central na comunidade, influenciando decisões políticas e econômicas.[76] Essa dinâmica contribuiu para a dificuldade de estabelecer instituições democráticas fortes e para a persistência de práticas clientelistas.[77]

Por fim, Freyre destaca a importância da compreensão interdisciplinar para analisar a sociedade brasileira.[78] Ele combina antropologia, sociologia, história e literatura para oferecer uma visão multifacetada das dinâmicas sociais.[79] Essa abordagem holística é essencial para entender como as estruturas patriarcais podem se manifestar em diferentes esferas, incluindo a tecnologia.[80]

A interdisciplinaridade permite uma análise mais profunda e complexa das questões sociais, evitando reducionismos e reconhecendo a interconexão entre diversos fatores.[81] No contexto atual, essa perspectiva é fundamental para investigar como a inteligência artificial não é apenas uma questão tecnológica, mas também social, cultural e política.[82]

3. Fundamentos da Inteligência Artificial

Para estabelecer uma conexão entre os conceitos de Gilberto Freyre e a inteligência artificial, é fundamental compreender o que é IA, como funciona o

[75] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[76] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[77] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[78] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[79] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[80] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[81] FREYRE, G. Casa-Grande & Senzala. Rio de Janeiro: Maia & Schmidt, 1933.
[82] SILVA, T. Inteligência Artificial e Sociedade: Reflexões Interdisciplinares. São Paulo: Editora da Universidade, 2020.

processo de treinamento com dados e o que significa viés no contexto de algoritmos e modelos de aprendizagem de máquina.

A inteligência artificial é um ramo da ciência da computação que busca criar sistemas capazes de realizar tarefas que normalmente exigiriam inteligência humana, como reconhecimento de padrões, tomada de decisão e compreensão de linguagem natural.[83] A IA engloba uma variedade de técnicas, incluindo aprendizado de máquina, aprendizado profundo e processamento de linguagem natural, que permitem aos sistemas aprender e melhorar a partir de dados sem serem explicitamente programados para cada tarefa específica.[84]

O desenvolvimento da IA remonta aos anos 1950, quando o termo foi cunhado por John McCarthy durante a Conferência de Dartmouth em 1956.[85] Desde então, a IA passou por diversas fases de entusiasmo e estagnação, conhecidas como "verões" e "invernos" da IA, refletindo os desafios técnicos e as expectativas sociais associadas à área.[86] Com os avanços em poder computacional, algoritmos e disponibilidade de dados, a IA contemporânea tem obtido sucessos significativos em múltiplos domínios.

Atualmente, a IA está presente em setores como saúde, finanças, transporte e educação. Aplicações incluem desde diagnósticos médicos assistidos por computador até veículos autônomos e sistemas de recomendação em plataformas de mídia social.[87] Esses desenvolvimentos não apenas transformam indústrias, mas também levantam questões éticas e sociais

---

[83] RUSSELL, S.; NORVIG, P. Artificial Intelligence: A Modern Approach. 3. ed. Upper Saddle River: Prentice Hall, 2016.
[84] GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep Learning. Cambridge: MIT Press, 2016.
[85] MCCARTHY, J. et al. A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. AI Magazine, v. 27, n. 4, p. 12-14, 2006.
[86] CREVIER, D. AI: The Tumultuous History of the Search for Artificial Intelligence. New York: Basic Books, 1993.
[87] JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. Science, v. 349, n. 6245, p. 255-260, 2015.

sobre o impacto da IA na sociedade, incluindo emprego, privacidade e equidade.[88]

O aprendizado de máquina é uma subárea da IA que se concentra no desenvolvimento de algoritmos que permitem aos computadores aprender a partir de dados.[89] Nesse contexto, os modelos são treinados utilizando grandes conjuntos de dados, ajustando parâmetros internos para minimizar a diferença entre suas previsões e os resultados reais.[90] Esse processo iterativo permite que o modelo identifique padrões e relações nos dados, melhorando seu desempenho em tarefas específicas.[91]

Existem diferentes paradigmas no aprendizado de máquina, incluindo aprendizado supervisionado, não supervisionado e por reforço. No aprendizado supervisionado, o modelo é treinado com pares de entrada e saída conhecidos, aprendendo a mapear entradas e saídas desejadas.[92] No aprendizado não supervisionado, o modelo busca encontrar estruturas ocultas ou padrões nos dados sem rótulos predefinidos.[93] Já o aprendizado por reforço envolve um agente que aprende a tomar decisões sequenciais para maximizar uma recompensa cumulativa, interagindo com um ambiente dinâmico.[94]

Aplicações práticas do aprendizado de máquina incluem reconhecimento de voz e imagem, processamento de linguagem natural, detecção de fraudes e

---

[88] FLORIDI, L. et al. AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, v. 28, n. 4, p. 689-707, 2018.
[89] MITCHELL, T. Machine Learning. New York: McGraw-Hill, 1997.
[90] ALPAYDIN, E. Introduction to Machine Learning. 4. ed. Cambridge: MIT Press, 2020.
[91] HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. The Elements of Statistical Learning. 2. ed. New York: Springer, 2009.
[92] AGGARWAL, C. C.; REDDY, C. K. Data Clustering: Algorithms and Applications. Boca Raton: CRC Press, 2013.
[93]
[94] SUTTON, R. S.; BARTO, A. G. Reinforcement Learning: An Introduction. 2. ed. Cambridge: MIT Press, 2018.

sistemas de recomendação.[95] A eficácia desses sistemas depende fortemente da qualidade dos dados de treinamento e da capacidade dos algoritmos de generalizar além dos dados observados, o que torna crucial a compreensão dos mecanismos subjacentes a esses modelos.[96]

O processo de treinamento envolve várias etapas críticas. Primeiro, é necessário coletar e preparar os dados, o que inclui a limpeza, a normalização e a segmentação em conjuntos de treinamento, validação e teste.[97] Em seguida, o modelo é definido e inicializado, e o algoritmo de aprendizado ajusta os parâmetros do modelo com base nos dados de treinamento.[98] Após cada iteração, o desempenho do modelo é avaliado usando o conjunto de validação, permitindo ajustes adicionais.[99] Finalmente, o modelo é testado com dados novos para verificar sua capacidade de generalização.[100]

A qualidade dos dados é um fator determinante para o sucesso do modelo. Dados incompletos, inconsistentes ou enviesados podem comprometer a capacidade do modelo de aprender padrões significativos.[101] Técnicas de pré-processamento, como imputação de valores ausentes, detecção de outliers e normalização, são essenciais para preparar os dados para o aprendizado.[102]

---

[95] DOMINGOS, P. The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World. New York: Basic Books, 2015.
[96] DOMINGOS, P. The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World. New York: Basic Books, 2015.
[97] KELLEHER, J. D.; NAMEE, B. M.; DOOLEY, A. Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies. Cambridge: MIT Press, 2015.
[98] KELLEHER, J. D.; NAMEE, B. M.; DOOLEY, A. Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies. Cambridge: MIT Press, 2015.
[99] MURPHY, K. P. Machine Learning: A Probabilistic Perspective. Cambridge: MIT Press, 2012.
[100] MURPHY, K. P. Machine Learning: A Probabilistic Perspective. Cambridge: MIT Press, 2012.
[101] PROVOST, F.; FAWCETT, T. Data science and its relationship to big data and data-driven decision making. Big Data, v. 1, n. 1, p. 51-59, 2013.
[102] PROVOST, F.; FAWCETT, T. Data science and its relationship to big data and data-driven decision making. Big Data, v. 1, n. 1, p. 51-59, 2013.

Além disso, o *overfitting* é um desafio comum, onde o modelo se ajusta tão bem aos dados de treinamento que perde a capacidade de generalizar para novos dados.[103] Estratégias como regularização, validação cruzada e seleção de características ajudam a mitigar esse problema, garantindo que o modelo mantenha um desempenho robusto em dados não vistos anteriormente.[104]

Um aspecto crítico no treinamento de modelos de IA é a questão do viés. O viés em IA refere-se a tendências ou preconceitos sistêmicos que levam a resultados injustos ou discriminatórios.[105] Esses vieses podem ser introduzidos de várias formas, incluindo dados históricos que refletem desigualdades sociais, amostras não representativas ou pressupostos inadequados nos algoritmos.[106] Como resultado, modelos treinados com dados tendenciosos podem perpetuar ou amplificar desigualdades existentes.[107]

O viés nos dados pode refletir desigualdades estruturais na sociedade. Por exemplo, se um conjunto de dados utilizado para treinar um modelo de reconhecimento facial contém predominantemente imagens de indivíduos de uma determinada etnia, o modelo pode apresentar desempenho inferior para outras etnias.[108] Isso destaca a importância de conjuntos de dados diversificados e representativos.

Além disso, o viés de confirmação pode ocorrer quando os pesquisadores interpretam os resultados de maneira a confirmar suas hipóteses prévias,

---

[103] GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep Learning. Cambridge: MIT Press, 2016.
[104] GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep Learning. Cambridge: MIT Press, 2016.
[105] BAROCAS, S.; SELBST, A. D. Big data's disparate impact. California Law Review, v. 104, n. 3, p. 671-732, 2016.
[106] OLTEANU, A. et al. Social data: Biases, methodological pitfalls, and ethical boundaries. Frontiers in Big Data, v. 2, p. 13, 2019.
[107] OLTEANU, A. et al. Social data: Biases, methodological pitfalls, and ethical boundaries. Frontiers in Big Data, v. 2, p. 13, 2019.
[108] BUOLAMWINI, J.; GEBRU, T. Gender shades: Intersectional accuracy disparities in commercial gender classification. In: Proceedings of the 1st Conference on Fairness, Accountability and Transparency. PMLR, 2018. p. 77-91.

negligenciando evidências contrárias.[109] Isso pode levar ao reforço de estereótipos e preconceitos existentes, afetando a objetividade e a imparcialidade dos modelos.[110]

Existem diferentes tipos de vieses em IA. O viés de seleção ocorre quando os dados de treinamento não representam adequadamente a população alvo, levando a modelos que funcionam bem para alguns grupos, mas não para outros.[111] O viés de confirmação pode surgir quando os desenvolvedores buscam ou interpretam dados de maneira a confirmar suas hipóteses pré-existentes.[112] Além disso, o viés algorítmico pode ocorrer devido a escolhas nos parâmetros ou na estrutura do modelo que favorecem determinados resultados.[113]

O viés histórico é outro tipo importante, onde modelos aprendem e perpetuam preconceitos presentes nos dados históricos, como discriminação de gênero ou racial.[114] Por exemplo, um algoritmo de recrutamento treinado em dados que refletem práticas discriminatórias pode continuar excluindo certos grupos demográficos.[115]

O viés de interação ocorre quando a interação dos usuários com o sistema influencia os resultados de forma enviesada.[116] Em plataformas de mídia

---

[109] NICKERSON, R. S. Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology, v. 2, n. 2, p. 175-220, 1998.
[110] NICKERSON, R. S. Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology, v. 2, n. 2, p. 175-220, 1998.
[111] ŽLIOBAITĖ, I. Measuring discrimination in algorithmic decision making. Data Mining and Knowledge Discovery, v. 31, n. 4, p. 1060-1089, 2017.
[112] NICKERSON, R. S. Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology, v. 2, n. 2, p. 175-220, 1998.
[113] FRIEDMAN, B.; NISSENBAUM, H. Bias in computer systems. ACM Transactions on Information Systems, v. 14, n. 3, p. 330-347, 1996.
[114] BAROCAS, S.; SELBST, A. D. Big data's disparate impact. California Law Review, v. 104, n. 3, p. 671-732, 2016.
[115] BAROCAS, S.; SELBST, A. D. Big data's disparate impact. California Law Review, v. 104, n. 3, p. 671-732, 2016.
[116] PARISER, E. The Filter Bubble: What the Internet is Hiding from You. New York: Penguin Press, 2011.

social, algoritmos de recomendação podem criar bolhas de filtro, expondo os usuários apenas a conteúdos que reforçam suas visões existentes.[117]

A mitigação de vieses é um desafio importante na IA. Técnicas para abordar esse problema incluem a coleta de dados mais representativos, o uso de algoritmos de aprendizado justo e a implementação de auditorias regulares nos modelos.[118] A transparência e a explicabilidade dos modelos também são fundamentais para identificar e corrigir vieses ocultos.[119]

Métodos como o re-amostragem ou o reponderação dos dados podem ajudar a equilibrar conjuntos de dados desbalanceados, reduzindo o viés de seleção.[120] Além disso, algoritmos podem ser ajustados para penalizar resultados injustos ou para impor restrições de equidade durante o treinamento.[121]

A participação multidisciplinar no desenvolvimento de sistemas de IA, incluindo especialistas em ética, sociologia e direito, pode contribuir para a identificação de potenciais vieses e para a elaboração de soluções que considerem o impacto social e ético da tecnologia.[122]

Outro conceito relevante é o de discriminação algorítmica, que ocorre quando sistemas de IA tomam decisões que prejudicam grupos específicos,

---

[117] ZHANG, B. H.; MITCHELL, M.; BAEZA-YATES, R. Fairness, accountability, and transparency in machine learning: Biases, methodologies, and challenges. AI Magazine, v. 39, n. 3, p. 55-66, 2018.
[118] GILPIN, L. H. et al. Explaining explanations: An overview of interpretability of machine learning. In: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA). IEEE, 2018. p. 80-89.
[119] GILPIN, L. H. et al. Explaining explanations: An overview of interpretability of machine learning. In: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA). IEEE, 2018. p. 80-89.
[120] KAMIRAN, F.; CALDERS, T. Data preprocessing techniques for classification without discrimination. Knowledge and Information Systems, v. 33, n. 1, p. 1-33, 2012.
[121] ZEMEL, R. S.; WU, Y.; SWERSKY, K.; PITASSI, T.; DWORK, C. Learning fair representations. In: Proceedings of the 30th International Conference on Machine Learning. Atlanta: PMLR, 2013. p. 325-333.
[122] SELBST, A. D. et al. Fairness and abstraction in sociotechnical systems. In: Proceedings of the Conference on Fairness, Accountability, and Transparency. 2019. p. 59-68.

muitas vezes refletindo preconceitos sociais.[123] Por exemplo, sistemas de reconhecimento facial têm apresentado taxas de erro mais altas para mulheres e pessoas negras, devido a conjuntos de dados desequilibrados e vieses nos algoritmos.[124] Isso evidencia como a IA pode perpetuar desigualdades se não for desenvolvida e implementada com cuidado.[125]

Casos de discriminação algorítmica incluem algoritmos de crédito que negam empréstimos a minorias ou sistemas de vigilância que monitoram desproporcionalmente certas comunidades.[126] Esses problemas destacam a necessidade de regulamentação e de práticas éticas no desenvolvimento e implantação de sistemas de IA.[127]

A transparência algorítmica e a possibilidade de auditoria são fundamentais para identificar e corrigir práticas discriminatórias.[128] Isso inclui não apenas a abertura do código-fonte, mas também a explicação clara dos critérios e dados utilizados para treinar e operar os modelos.[129]

---

[123] NOBLE, S. U. Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press, 2018.
[124] BUOLAMWINI, J.; GEBRU, T. Gender shades: Intersectional accuracy disparities in commercial gender classification. In: Proceedings of the 1st Conference on Fairness, Accountability and Transparency. PMLR, 2018. p. 77-91.
[125] BUOLAMWINI, J.; GEBRU, T. Gender shades: Intersectional accuracy disparities in commercial gender classification. In: Proceedings of the 1st Conference on Fairness, Accountability and Transparency. PMLR, 2018. p. 77-91.
[126] ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica, 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Acesso em: 10 out. 2023.
[127] ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica, 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Acesso em: 10 out. 2023.
[128] PASQUALE, F. The Black Box Society: The Secret Algorithms That Control Money and Information. Cambridge: Harvard University Press, 2015.
[129] PASQUALE, F. The Black Box Society: The Secret Algorithms That Control Money and Information. Cambridge: Harvard University Press, 2015.

A ética na IA é um campo emergente que busca abordar questões como responsabilidade, justiça, privacidade e transparência.[130] Diretrizes éticas visam orientar o desenvolvimento de sistemas de IA que respeitem os direitos humanos e promovam benefícios sociais. Organizações internacionais e governos têm elaborado princípios e regulamentos para garantir o uso responsável da IA.[131]

Princípios como o de não maleficência, que busca evitar causar danos, e o de justiça, que promove a equidade e evita a discriminação, são centrais nesses debates.[132] A responsabilidade também é enfatizada, assegurando que haja prestação de contas por parte dos desenvolvedores e operadores de sistemas de IA.[133]

Iniciativas como a Parceria Global em Inteligência Artificial (GPAI) e as diretrizes da Comissão Europeia refletem esforços internacionais para estabelecer padrões éticos e regulatórios na IA.[134] Essas diretrizes visam equilibrar a inovação tecnológica com a proteção dos valores humanos fundamentais.[135]

O aprendizado profundo é outra subárea da IA que utiliza redes neurais artificiais com múltiplas camadas para modelar padrões complexos em dados.[136] Embora essa técnica tenha alcançado avanços significativos em áreas

---

[130] FLORIDI, L. et al. AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, v. 28, n. 4, p. 689-707, 2018.
[131] JOBIN, A.; IENCA, M.; VAYENA, E. The global landscape of AI ethics guidelines. Nature Machine Intelligence, v. 1, n. 9, p. 389-399, 2019.
[132] FLORIDI, L. et al. AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, v. 28, n. 4, p. 689-707, 2018.
[133] FLORIDI, L. et al. AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, v. 28, n. 4, p. 689-707, 2018.
[134] GPAI. Global Partnership on Artificial Intelligence. 2020. Disponível em: https://gpai.ai/. Acesso em: 10 out. 2023.
[135] GPAI. Global Partnership on Artificial Intelligence. 2020. Disponível em: https://gpai.ai/. Acesso em: 10 out. 2023.
[136] LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. Nature, v. 521, n. 7553, p. 436-444, 2015.

como visão computacional e processamento de linguagem natural, também apresenta desafios em termos de interpretabilidade e controle de vieses.[137] Devido à complexidade dessas redes, pode ser difícil entender como elas chegam a determinadas decisões, o que dificulta a identificação de possíveis vieses.[138]

As redes neurais convolucionais (CNNs) são amplamente utilizadas em reconhecimento de imagens, enquanto as redes neurais recorrentes (RNNs) são aplicadas em processamento de linguagem natural e reconhecimento de voz.[139] Essas arquiteturas permitem que modelos de aprendizado profundo capturem dependências complexas nos dados.[140]

No entanto, a opacidade dos modelos de aprendizado profundo, frequentemente referida como "caixa-preta", levanta preocupações sobre a confiabilidade e a responsabilidade nas decisões automatizadas.[141] Isso é particularmente crítico em aplicações de alto risco, como diagnóstico médico ou decisões judiciais, onde a falta de explicabilidade pode impedir a detecção de erros ou vieses.[142]

Finalmente, é importante mencionar o conceito de inteligência artificial explicável (XAI), que busca desenvolver modelos de IA cujas decisões possam ser facilmente compreendidas por humanos.[143] A XAI é crucial para construir

---

[137] RUDIN, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, v. 1, n. 5, p. 206-215, 2019.
[138] RUDIN, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, v. 1, n. 5, p. 206-215, 2019.
[139] SCHMIDHUBER, J. Deep learning in neural networks: An overview. Neural Networks, v. 61, p. 85-117, 2015.
[140] SCHMIDHUBER, J. Deep learning in neural networks: An overview. Neural Networks, v. 61, p. 85-117, 2015.
[141] DOSHI-VELEZ, F.; KIM, B. Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608, 2017.
[142] DOSHI-VELEZ, F.; KIM, B. Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608, 2017.
[143] ADADI, A.; BERRADA, M. Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). IEEE Access, v. 6, p. 52138-52160, 2018.

confiança nos sistemas de IA e para garantir que decisões automatizadas possam ser auditadas e justificadas, especialmente em contextos sensíveis como saúde, justiça e finanças.[144]

Métodos de XAI incluem técnicas como LIME (Local Interpretable Model-agnostic Explanations) e SHAP (SHapley Additive exPlanations), que oferecem explicações locais sobre as previsões de modelos complexos.[145] Além disso, o desenvolvimento de modelos interpretable-by-design, como árvores de decisão e modelos lineares, é incentivado em situações onde a explicabilidade é essencial.[146]

A regulamentação também está evoluindo para exigir maior transparência em sistemas de IA. O Regulamento Geral sobre a Proteção de Dados (GDPR) da União Europeia inclui o direito a uma explicação sobre decisões automatizadas, refletindo a importância da XAI no contexto legal e ético.[147]

Compreender esses fundamentos da inteligência artificial é essencial para analisar como as estruturas sociais e culturais, incluindo as patriarcais descritas por Gilberto Freyre, podem influenciar o desenvolvimento e a aplicação da IA. A interação entre tecnologia e sociedade é complexa, e reconhecer os potenciais vieses e desigualdades é o primeiro passo para promover sistemas de IA mais justos e inclusivos.

---

[144] ADADI, A.; BERRADA, M. Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). IEEE Access, v. 6, p. 52138-52160, 2018.
[145] RIBEIRO, M. T.; SINGH, S.; GUESTRIN, C. "Why should I trust you?": Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016. p. 1135-1144.
[146] RUDIN, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, v. 1, n. 5, p. 206-215, 2019.
[147] GOODMAN, B.; FLAXMAN, S. European Union regulations on algorithmic decision-making and a "right to explanation". AI Magazine, v. 38, n. 3, p. 50-57, 2017.

4. Viés e Discriminação Algorítmica

A expansão da inteligência artificial e do aprendizado de máquina em diversas áreas tem trazido à tona preocupações significativas relacionadas ao viés e à discriminação algorítmica. Esses fenômenos ocorrem quando sistemas automatizados tomam decisões que prejudicam determinados grupos sociais, muitas vezes refletindo e amplificando preconceitos existentes na sociedade.[148]

Uma das principais fontes de viés é a falta de representatividade nos dados de treinamento. Dados que não refletem a diversidade da população podem levar a modelos com desempenho desigual entre diferentes grupos.[149] Por exemplo, sistemas de reconhecimento de fala treinados predominantemente com vozes masculinas podem apresentar dificuldades em compreender vozes femininas ou sotaques diversos.[150] Além disso, feedback loops podem exacerbar o viés, onde as decisões do algoritmo influenciam os dados futuros, reforçando tendências existentes.[151]

Esse tipo de viés pode perpetuar estereótipos e discriminação institucional. Algoritmos de recrutamento podem descartar currículos com nomes associados a minorias étnicas, seguindo padrões históricos de contratação.[152] No setor financeiro, modelos de risco de crédito podem negar empréstimos a indivíduos de certas áreas geográficas ou com histórico

---

[148] NOBLE, S. U. Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press, 2018.
[149] ZLIOBAITĖ, I. Measuring discrimination in algorithmic decision making. Data Mining and Knowledge Discovery, v. 31, n. 4, p. 1060-1089, 2017.
[150] TATMAN, R. Gender and dialect bias in YouTube's automatic captions. In: Proceedings of the First ACL Workshop on Ethics in Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2017. p. 53-59.
[151] ENSIGN, D.; FRIEDLER, S. A.; NEVILLE, S.; SCHEIDEGGER, C.; VENKATASUBRAMANIAN, S. Runaway feedback loops in predictive policing. In: Proceedings of the 1st Conference on Fairness, Accountability and Transparency. PMLR, 2018. p. 160-171.
[152] BERTRAND, M.; MULLAINATHAN, S. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. American Economic Review, v. 94, n. 4, p. 991-1013, 2004.

financeiro limitado, correlacionando-se com status socioeconômico ou origem étnica.[153] No contexto judicial, ferramentas de avaliação de risco como o COMPAS foram criticadas por atribuir pontuações mais altas de reincidência a indivíduos negros em comparação com brancos, mesmo quando controlados fatores relevantes.[154]

A falta de transparência em sistemas de IA, frequentemente operando como "caixas-pretas", dificulta a identificação e correção de vieses.[155] Isso levanta questões sobre responsabilidade e prestação de contas, já que indivíduos afetados por decisões algorítmicas discriminatórias têm poucas opções de recurso ou contestação.[156]

Para abordar o viés e a discriminação algorítmica, é essencial implementar práticas de justiça algorítmica. Isso inclui a adoção de técnicas de mitigação de viés, como reamostragem de dados, ajustes nos algoritmos para penalizar desigualdades e o monitoramento contínuo dos resultados.[157] Ferramentas como o aprendizado de máquina justo buscam incorporar critérios de equidade diretamente nos algoritmos, garantindo desempenho consistente entre

---

[153] HURLEY, M.; ADEBAYO, J. Credit scoring in the era of big data. Yale Journal of Law & Technology, v. 18, p. 148-216, 2016.
[154] ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica, 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Acesso em: 10 out. 2023.
[155] PASQUALE, F. The Black Box Society: The Secret Algorithms That Control Money and Information. Cambridge: Harvard University Press, 2015.
[156] CITRON, D. K.; PASQUALE, F. The scored society: Due process for automated predictions. Washington Law Review, v. 89, n. 1, p. 1-33, 2014.
[157] ZEMEL, R. S.; WU, Y.; SWERSKY, K.; PITASSI, T.; DWORK, C. Learning fair representations. In: Proceedings of the 30th International Conference on Machine Learning. Atlanta: PMLR, 2013. p. 325-333.

diferentes grupos.[158] A inclusão de perspectivas diversas nas equipes de desenvolvimento também é crucial.[159]

Regulamentações e diretrizes éticas desempenham um papel importante na promoção de sistemas de IA justos e equitativos. A União Europeia, por exemplo, propôs o Regulamento de Inteligência Artificial, que estabelece requisitos para sistemas de IA de alto risco, incluindo avaliações de conformidade e obrigações de transparência.[160] Organizações como a IEEE e a UNESCO também têm elaborado diretrizes éticas para orientar o desenvolvimento e a implementação de IA.[161] No entanto, a eficácia dessas regulamentações depende de sua aplicação prática e da cooperação entre diferentes stakeholders, incluindo governos, setor privado e sociedade civil.[162]

No contexto brasileiro, a discussão sobre viés e discriminação algorítmica é particularmente relevante devido às desigualdades históricas relacionadas a gênero, raça e classe social.[163] Casos como o uso de algoritmos para determinar elegibilidade em programas sociais ou para identificar áreas prioritárias em políticas de segurança pública podem impactar desproporcionalmente comunidades vulneráveis.[164] Além disso, a falta de diversidade nas indústrias de

---

[158] ZEMEL, R. S.; WU, Y.; SWERSKY, K.; PITASSI, T.; DWORK, C. Learning fair representations. In: Proceedings of the 30th International Conference on Machine Learning. Atlanta: PMLR, 2013. p. 325-333.
[159] CRAWFORD, K.; CALO, R. There is a blind spot in AI research. Nature, v. 538, n. 7625, p. 311-313, 2016.
[160] UNIÃO EUROPEIA. Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Bruxelas, 2021.
[161] UNESCO. Recommendation on the Ethics of Artificial Intelligence. Paris: UNESCO, 2021; INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS (IEEE). Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems. New York: IEEE, 2019.
[162] JOBIN, A.; IENCA, M.; VAYENA, E. The global landscape of AI ethics guidelines. Nature Machine Intelligence, v. 1, n. 9, p. 389-399, 2019.
[163] MEDEIROS, M.; ALBERTI, V. R. (Org.). Desigualdades Sociais, Violências e Racismo no Brasil. Rio de Janeiro: Editora FGV, 2020.
[164] CRUZ, F. A.; SILVA, R. M. Inteligência artificial e discriminação algorítmica: desafios para a regulação no Brasil. Revista de Direito, Estado e Telecomunicações, v. 12, n. 1, p. 1-22, 2020.

tecnologia no Brasil pode limitar a capacidade de identificar e abordar vieses nos sistemas.[165]

Iniciativas como a Lei Geral de Proteção de Dados (LGPD) e debates sobre a criação de um marco regulatório específico para IA refletem a crescente preocupação com essas questões.[166] A participação ativa de acadêmicos, organizações não governamentais e comunidades afetadas é essencial para a elaboração de políticas eficazes e inclusivas.

5. A Reprodução de Estruturas Patriarcais pela IA

Ao ser alimentada com dados históricos e contemporâneos, a inteligência artificial reflete o mundo em que esses dados foram gerados. Se os dados representam uma sociedade com profundas desigualdades de gênero, como a brasileira descrita por Gilberto Freyre, a IA pode reforçar esses padrões, perpetuando a dominação masculina e as relações de poder associadas ao patriarcalismo.[167]

Gilberto Freyre enfatizou que o patriarcado brasileiro não era apenas uma estrutura familiar, mas uma organização social que permeava todas as esferas da vida, desde a economia até a política.[168] Ao considerar que os dados utilizados para treinar sistemas de IA são extraídos de contextos históricos e sociais marcados por essa dominação masculina, é inevitável que os algoritmos reflitam e reforcem essas estruturas.[169]

---

[165] SILVA, M. L.; FIGUEIREDO, D. R. Diversidade na área de computação no Brasil: uma análise de gênero e raça. Revista Brasileira de Informática na Educação, v. 29, p. 529-551, 2021.
[166] BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Diário Oficial da União, Brasília, DF, 15 ago. 2018.; CÂMARA DOS DEPUTADOS. Projeto de Lei 21/2020. Dispõe sobre o uso da inteligência artificial no Brasil. Brasília, DF, 2020.
[167] NOBLE, S. U. Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press, 2018.
[168] FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).
[169] CRAWFORD, K.; CALO, R. There is a blind spot in AI research. Nature, v. 538, n. 7625, p. 311-313, 2016.

Além disso, a IA, ao operar com base em grandes volumes de dados, muitas vezes carece da capacidade de contextualizar ou questionar os padrões existentes. Isso significa que, sem intervenções conscientes, os sistemas de IA podem perpetuar práticas discriminatórias que já deveriam ter sido superadas. Assim como o patriarcado descrito por Freyre se adaptou e persistiu ao longo do tempo, a IA pode atuar como um novo meio pelo qual essas estruturas opressivas são mantidas na sociedade contemporânea.[170]

Essa perpetuação ocorre porque os algoritmos de IA aprendem a partir de padrões existentes nos dados de treinamento. Se esses dados contêm vieses de gênero, raça ou classe, os modelos resultantes tendem a reproduzir essas desigualdades.[171] Portanto, a IA não é uma entidade neutra; ela carrega consigo as estruturas sociais e culturais inerentes aos dados que a alimentam.

Freyre destacou que as relações de poder no Brasil foram moldadas por uma hierarquia rígida, na qual o patriarca não apenas controlava a família, mas também influenciava a comunidade e as instituições.[172] Da mesma forma, os sistemas de IA podem assumir um papel de autoridade, tomando decisões que afetam a vida das pessoas sem transparência ou responsabilidade clara.[173] Isso reflete a centralização do poder nas mãos de poucos, característica do patriarcado colonial.

A falta de diversidade nas equipes de desenvolvimento de IA contribui para a falta de consciência sobre os vieses incorporados nos algoritmos.[174] Essa

---

[170] D'IGNACIO, C.; KLEIN, L. Data Feminism. Cambridge: MIT Press, 2020.
[171] BAROCAS, S.; SELBST, A. D. Big data's disparate impact. California Law Review, v. 104, n. 3, p. 671-732, 2016.
[172] FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).
[173] O'NEIL, C. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown, 2016.
[174] HICKS, M. Programmed Inequality: How Britain Discarded Women Technologists and Lost Its Edge in Computing. Cambridge: MIT Press, 2017.

homogeneidade é um espelho da sociedade patriarcal descrita por Freyre, onde a voz e a experiência de grupos marginalizados são excluídas dos processos decisórios. Sem a inclusão de perspectivas diversas, a IA continuará a reforçar as estruturas de poder existentes.[175]

Por exemplo, sistemas de IA usados em recrutamento têm sido criticados por favorecerem candidatos masculinos em detrimento das mulheres.[176] Isso ocorre porque os algoritmos são treinados com base em dados históricos de contratação, que frequentemente refletem o preconceito de gênero existente nas indústrias dominadas por homens. Dessa forma, a IA "aprende" que candidatos homens são mais adequados, reproduzindo a desigualdade no processo de seleção.

Essa situação reflete o que Freyre descreveu sobre a transmissão intergeracional de privilégios dentro da estrutura patriarcal.[177] Assim como os filhos homens herdavam o poder e os recursos do patriarca, excluindo mulheres e outros grupos, a IA perpetua essa transmissão de privilégios ao favorecer candidatos que se encaixam no perfil historicamente dominante. Isso impede a mobilidade social e a diversificação dos espaços profissionais.[178]

A reprodução de desigualdades no mercado de trabalho tem consequências amplas para a sociedade. A falta de diversidade não apenas afeta a equidade nos processos de contratação, mas também limita a inovação e a capacidade das organizações de atender às necessidades de uma população

[175] CRAWFORD, K.; CALO, R. There is a blind spot in AI research. Nature, v. 538, n. 7625, p. 311-313, 2016.
[176] BOLUKBASI, T.; CHANG, K.-W.; ZOU, J.; SALIGRAMA, V.; KALAI, A. Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In: Advances in Neural Information Processing Systems 29 (NIPS 2016). [S.l.: s.n.], 2016. p. 4349-4357.
[177] FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).
[178] D'IGNACIO, C.; KLEIN, L. Data Feminism. Cambridge: MIT Press, 2020.

diversa.[179] Assim, a IA, ao reforçar padrões discriminatórios, contribui para a manutenção das estruturas patriarcais identificadas por Freyre.

Freyre também apontou como o patriarcado estabelecia normas culturais que legitimavam a desigualdade, naturalizando a subordinação de mulheres e outros grupos.[180] Da mesma forma, a IA, ao operar de forma opaca e sem questionamento, pode legitimar decisões discriminatórias como se fossem objetivas ou imparciais.[181] Isso dificulta a identificação e o combate às injustiças, uma vez que a autoridade da tecnologia é raramente contestada. A centralização do poder decisório na IA pode levar à diminuição da responsabilidade humana, criando uma situação em que ninguém se sente responsável pelos resultados prejudiciais.[182]

Além de perpetuar a desigualdade de gênero, a IA também reproduz outras hierarquias, como as de raça e classe.[183] No contexto brasileiro, onde a dominação patriarcal se entrelaça com dinâmicas raciais e socioeconômicas, isso se torna particularmente preocupante. Ao serem treinados com dados que refletem séculos de exclusão e desigualdade, os sistemas de IA reforçam essas mesmas estruturas de poder.

Freyre destacou a complexidade das relações raciais no Brasil, incluindo a influência do patriarcado na manutenção das desigualdades entre brancos, negros e indígenas.[184] A miscigenação, embora presente, não eliminou as

---

[179] PAGE, S. E. The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies. Princeton: Princeton University Press, 2007.
[180] FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).
[181] PASQUALE, F. The Black Box Society: The Secret Algorithms That Control Money and Information. Cambridge: Harvard University Press, 2015.
[182] JOHNSON, D. G. Technology with no human responsibility? Journal of Business Ethics, v. 127, n. 4, p. 707-715, 2015.
[183] EUBANKS, V. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin's Press, 2018.
[184] FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).

hierarquias raciais, mas muitas vezes as mascarou sob o mito da democracia racial.[185] A IA, ao não reconhecer essas nuances, pode contribuir para a invisibilização e marginalização de grupos étnicos.[186]

Sistemas de reconhecimento facial, por exemplo, têm demonstrado taxas de erro significativamente maiores para pessoas negras, devido à sub-representação nos dados de treinamento.[187] Isso não só evidencia a reprodução de vieses raciais, mas também levanta preocupações sobre segurança e privacidade, especialmente quando essas tecnologias são utilizadas por órgãos de segurança pública, potencialmente reforçando práticas discriminatórias.[188]

Freyre observou como a concentração de riqueza e poder nas mãos do patriarca limitava o acesso de outros grupos a recursos e oportunidades.[189] Da mesma forma, a IA pode atuar como um guardião tecnológico que impede o acesso de indivíduos marginalizados a serviços essenciais. Isso reforça as barreiras socioeconômicas e dificulta a redução das desigualdades.[190]

A utilização de IA em decisões financeiras sem transparência ou possibilidade de contestação reforça a sensação de injustiça e impotência entre os afetados. Essa falta de agência é similar à subordinação enfrentada por grupos marginalizados no patriarcado colonial, onde as decisões eram impostas sem consideração pelas necessidades ou direitos dos indivíduos.[191]

---

[185] MUNANGA, K. Rediscutindo a Mestiçagem no Brasil: Identidade Nacional versus Identidade Negra. Petrópolis: Vozes, 1999.
[186] NOBLE, S. U. Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press, 2018.
[187] BUOLAMWINI, J.; GEBRU, T. Gender shades: intersectional accuracy disparities in commercial gender classification. In: Conference on Fairness, Accountability and Transparency, 2018. Proceedings... p.77-91.
[188] LEHMANN, J. F. et al. Facial recognition technologies in the wild: A call for a federal office. Science, v. 365, n. 6455, p. 128-130, 2019.
[189] FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).
[190] EUBANKS, V. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin's Press, 2018.
[191] FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).

No sistema judicial, o uso de IA para prever a probabilidade de reincidência criminal tem levantado preocupações sobre discriminação racial.[192] Modelos treinados com dados históricos podem atribuir maior risco a indivíduos de determinadas etnias, influenciando decisões de sentenciamento e liberdade condicional. Isso reflete a perpetuação de estereótipos e preconceitos arraigados na sociedade.

A dependência em sistemas automatizados pode levar à desumanização dos indivíduos, reduzindo-os a estatísticas ou perfis de risco. Isso contrasta com a necessidade de considerar contextos individuais e circunstâncias atenuantes, princípios fundamentais para um sistema de justiça equitativo. A falta de empatia e compreensão humana é agravada quando decisões críticas são delegadas a algoritmos.[193]

A reprodução dessas estruturas pela IA não é apenas um problema técnico, mas também ético e social.[194] A falta de diversidade nas equipes que desenvolvem esses sistemas contribui para a falta de consciência sobre os vieses incorporados.[195] Sem a inclusão de perspectivas diversas, os sistemas de IA continuarão a refletir e amplificar as desigualdades existentes.

Freyre enfatizou a importância de reconhecer as múltiplas influências culturais e sociais na formação da sociedade brasileira.[196] Da mesma forma, é crucial que o desenvolvimento de IA inclua diversas vozes e experiências para

---

[192] ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica, 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Acesso em: 10 out. 2023.
[193] BARRETT, L. F. How Emotions Are Made: The Secret Life of the Brain. New York: Houghton Mifflin Harcourt, 2017.
[194] O'NEIL, C. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown, 2016.
[195] HICKS, M. Programmed Inequality: How Britain Discarded Women Technologists and Lost Its Edge in Computing. Cambridge: MIT Press, 2017.
[196] FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).

evitar a reprodução de estruturas opressivas. A inclusão de mulheres, pessoas de diferentes origens étnicas e socioeconômicas pode enriquecer o processo de desenvolvimento e resultar em sistemas mais justos.[197]

Ao integrar os conceitos de Freyre com a análise crítica da IA, podemos desenvolver estratégias que não apenas evitem a reprodução de injustiças, mas que também utilizem a tecnologia como meio de promover a equidade. Isso exige uma abordagem interdisciplinar e reflexiva, reconhecendo que a tecnologia é moldada pelas escolhas humanas e pelos contextos sociais em que é desenvolvida.

A construção de sistemas de IA justos e inclusivos é um esforço contínuo que requer vigilância, educação e engajamento de todos os setores da sociedade. Somente assim poderemos aproveitar plenamente os benefícios da IA enquanto mitigamos seus riscos e desafios.

6. Patriarcalismo e Oligopólios Tecnológicos: A Concentração de Poder na Era dos Dados

Observa-se um paralelo entre os conceitos de Freyre e a concentração de poder nas grandes corporações de tecnologia, conhecidas como "Big Techs". Empresas como Google, Amazon, Facebook, Apple e Microsoft acumulam vasto controle sobre dados e informações, influenciando significativamente processos de governança e decisões sociais. Esse fenômeno é frequentemente denominado "colonialismo tecnológico" ou "oligopólio tecnológico".

O "colonialismo tecnológico" refere-se à maneira como essas corporações exercem influência global, muitas vezes sem considerar as

[197] D'IGNACIO, C.; KLEIN, L. Data Feminism. Cambridge: MIT Press, 2020.

particularidades culturais, sociais e políticas dos países onde operam.[198] Ao controlarem plataformas digitais amplamente utilizadas, estabelecem padrões de comportamento, consumo e comunicação, impondo uma hegemonia cultural e tecnológica.[199] Essa dominação tecnológica resulta na disseminação de valores e práticas que não refletem a diversidade das sociedades, semelhante à imposição de normas pelo patriarca nas comunidades tradicionais.

Adicionalmente, essas empresas coletam e analisam enormes quantidades de dados pessoais, possibilitando a criação de perfis detalhados de usuários e a predição de comportamentos.[200] Essa capacidade é utilizada para direcionar conteúdos, anúncios e até mesmo influenciar processos eleitorais.[201]

A ascensão da inteligência artificial amplifica essa dinâmica de poder. Tecnologias de IA são empregadas em sistemas de vigilância, algoritmos de recomendação e estratégias de manipulação de informação. Essas ferramentas tornam-se instrumentos para ampliar o alcance e a profundidade da influência das Big Techs na sociedade contemporânea, afetando desde escolhas individuais até processos políticos de grande escala.

Sistemas de IA permitem o processamento e a interpretação de dados em escala massiva, proporcionando insights profundos sobre o comportamento humano. Tecnologias como reconhecimento facial, análise de sentimentos e

[198] COULDRY, Nick; MEJIAS, Ulises A. The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism. Stanford: Stanford University Press, 2019.
[199] COULDRY, Nick; MEJIAS, Ulises A. The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism. Stanford: Stanford University Press, 2019.
[200] ZUBOFF, S. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs, 2019.
[201] ZUBOFF, S. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs, 2019.

algoritmos de personalização moldam significativamente as experiências individuais e coletivas no mundo digital e físico.[202]

Quando utilizados sem supervisão ou regulamentação adequadas, esses sistemas podem ser empregados para monitorar populações, suprimir dissidências e manipular opiniões públicas.[203] A opacidade nos processos algorítmicos e a ausência de mecanismos de responsabilização dificultam a identificação e correção de vieses ou abusos, perpetuando desigualdades e ameaçando a integridade dos processos democráticos.[204] Essa falta de transparência reforça a assimetria de poder entre as corporações e os indivíduos, semelhante à relação entre o patriarca e os membros subalternos de sua comunidade.

Essa concentração de poder tecnológico e informacional configura uma nova forma de patriarcalismo, na qual a autoridade é centralizada em entidades corporativas que exercem influência abrangente sobre a esfera pública e privada. As Big Techs atuam como atores dominantes, moldando decisões políticas e sociais de acordo com seus interesses corporativos, muitas vezes sem a participação ou o consentimento explícito da sociedade.

Essas empresas possuem a capacidade de determinar quais informações são disseminadas ou suprimidas, influenciando a formação de opiniões e decisões individuais e coletivas.[205] Por meio de algoritmos cuja lógica interna não é transparente ao público, controlam o fluxo de notícias e conteúdos nas plataformas digitais, podendo priorizar determinadas narrativas e silenciar

---

[202] HARARI, Y. N. 21 Lições para o Século 21. São Paulo: Companhia das Letras, 2018.
[203] O'NEIL, C. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown, 2016.
[204] O'NEIL, C. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown, 2016.
[205] PARISER, E. The Filter Bubble: What the Internet is Hiding from You. New York: Penguin Press, 2011.

outras.[206] Isso afeta diretamente o discurso público e a qualidade da democracia, já que a diversidade de vozes e o acesso a informações imparciais são essenciais para uma participação cidadã informada.[207]

A homogeneidade nas lideranças dessas corporações também contribui para uma visão de mundo limitada e enviesada, que não representa a pluralidade cultural e social da população global.[208] A predominância de perspectivas de um grupo restrito pode resultar na perpetuação de estereótipos e preconceitos nos sistemas de IA, reforçando desigualdades existentes e criando novas formas de exclusão. Essa falta de diversidade interna reflete a estrutura patriarcal, onde a concentração de poder em um grupo homogêneo impede a inclusão de diferentes vozes e experiências.

Para enfrentar essa nova forma de concentração de poder, é imprescindível promover mecanismos de governança que assegurem a transparência, a responsabilidade e a participação democrática no desenvolvimento e uso da IA. A regulamentação dessas tecnologias torna-se essencial para equilibrar os interesses corporativos com o bem-estar social e os direitos individuais. Políticas regulatórias e marcos éticos podem ser implementados para limitar o poder dessas corporações e garantir que a tecnologia seja utilizada de maneira que beneficie a sociedade como um todo.

A comparação entre o patriarcalismo freyriano e os oligopólios tecnológicos atuais evidencia a persistência de estruturas de poder centralizadas que podem limitar o desenvolvimento de sociedades mais justas

---

[206] PARISER, E. The Filter Bubble: What the Internet is Hiding from You. New York: Penguin Press, 2011.
[207] PARISER, E. The Filter Bubble: What the Internet is Hiding from You. New York: Penguin Press, 2011.
[208] NOBLE, S. U. Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press, 2018.

e democráticas. Assim como Freyre analisou criticamente o papel do patriarca na sociedade brasileira, é essencial examinar e questionar o poder das Big Techs e o impacto da IA nos processos sociais e políticos. A identificação dessas dinâmicas permite a elaboração de estratégias para desconstruir estruturas opressivas e promover a democratização da tecnologia.

7. Casos de Reprodução de Vieses pela IA

A reprodução de vieses em sistemas de inteligência artificial (IA) não é apenas uma possibilidade teórica, mas uma realidade evidenciada por diversos casos práticos. Esses incidentes ilustram como a IA, ao ser desenvolvida e implementada sem uma consideração cuidadosa dos vieses históricos e culturais, pode perpetuar e amplificar discriminações de gênero e raça.

7.1. O Algoritmo de Recrutamento da Amazon

Em 2018, a Amazon descontinuou um sistema de recrutamento baseado em IA após descobrir que o algoritmo discriminava candidatas mulheres.[209] O sistema foi treinado utilizando currículos submetidos à empresa ao longo de dez anos, período em que a maioria dos candidatos contratados eram homens, refletindo a predominância masculina no setor de tecnologia.[210] Como resultado, o algoritmo aprendeu a favorecer candidatos masculinos, penalizando palavras-chave associadas ao gênero feminino.[211]

---

[209] DASTIN, J. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters, 2018. Disponível em: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G . Acesso em: 10 out. 2023.
[210] DASTIN, J. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters, 2018. Disponível em: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G . Acesso em: 10 out. 2023.
[211] DASTIN, J. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters, 2018. Disponível em: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G . Acesso em: 10 out. 2023.

O algoritmo de recrutamento da Amazon foi concebido para automatizar a triagem de currículos, acelerando o processo de contratação e identificando os candidatos mais qualificados.[212] A ideia era que o sistema aprendesse a partir dos currículos enviados à empresa ao longo de uma década, utilizando técnicas de aprendizado de máquina para reconhecer padrões que levaram a contratações bem-sucedidas no passado.[213]

No entanto, devido ao fato de que a maioria dos currículos apresentados e das contratações efetivadas pela Amazon nesse período eram de homens, o sistema aprendeu a associar características masculinas a candidatos ideais.[214] Termos e experiências que sinalizavam gênero feminino eram penalizados, levando a uma subvalorização de candidatas mulheres, independentemente de suas qualificações profissionais.[215]

Essa situação evidencia como dados históricos carregados de vieses podem levar a resultados discriminatórios quando utilizados sem a devida crítica e ajuste em sistemas de IA.[216] O algoritmo não discriminou por intenção própria, mas sim refletiu os padrões presentes nos dados de treinamento, perpetuando desigualdades de gênero existentes no setor de tecnologia.[217]

---

[212] DASTIN, J. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters, 2018. Disponível em: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G . Acesso em: 10 out. 2023.
[213] DASTIN, J. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters, 2018. Disponível em: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G . Acesso em: 10 out. 2023.
[214] DASTIN, J. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters, 2018. Disponível em: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G . Acesso em: 10 out. 2023.
[215] DASTIN, J. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters, 2018. Disponível em: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G . Acesso em: 10 out. 2023.
[216] BOLUKBASI, T.; CHANG, K.-W.; ZOU, J.; SALIGRAMA, V.; KALAI, A. Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In: Advances in Neural Information Processing Systems 29 (NIPS 2016). [S.l.: s.n.], 2016. p. 4349-4357.
[217] BOLUKBASI, T.; CHANG, K.-W.; ZOU, J.; SALIGRAMA, V.; KALAI, A. Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In: Advances in Neural Information Processing Systems 29 (NIPS 2016). [S.l.: s.n.], 2016. p. 4349-4357.

Após a descoberta do viés, a Amazon descontinuou o uso do sistema, reconhecendo as limitações e riscos associados.[218] Este caso serviu como um alerta para a indústria sobre a necessidade de abordar ativamente os vieses nos sistemas de IA, implementando práticas de desenvolvimento éticas e inclusivas, como a diversificação dos conjuntos de dados e a realização de auditorias regulares de viés.[219]

7.2. O Algoritmo COMPAS no Sistema Judicial

O algoritmo COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*) é um software utilizado nos Estados Unidos para prever a probabilidade de reincidência criminal.[220] Uma investigação realizada pelo ProPublica revelou que o COMPAS apresentava vieses raciais significativos, atribuindo pontuações de risco mais altas a réus negros em comparação com réus brancos com históricos semelhantes.[221]

O COMPAS foi amplamente adotado nos tribunais dos Estados Unidos para auxiliar juízes na determinação de sentenças e decisões sobre liberdade condicional, baseando-se em uma pontuação que indica o risco de reincidência

---

[218] DASTIN, J. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters, 2018. Disponível em: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G . Acesso em: 10 out. 2023.
[219] DASTIN, J. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters, 2018. Disponível em: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G . Acesso em: 10 out. 2023.
[220] ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica, 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing . Acesso em: 10 out. 2023.
[221] ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica, 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing . Acesso em: 10 out. 2023.

criminal.[222] A promessa era de que a utilização de algoritmos poderia trazer maior objetividade e eficiência ao sistema judicial.[223]

No entanto, a investigação revelou que o COMPAS apresentava um viés significativo contra réus negros.[224] O estudo mostrou que o algoritmo tinha uma taxa de falsos positivos mais alta para negros, classificando-os erroneamente como de alto risco de reincidência, enquanto atribuía uma taxa de falsos negativos mais alta para réus brancos, subestimando seu risco real.[225]

Essa disparidade ocorre devido a fatores como a consideração de histórico criminal e ambiente social, que estão profundamente enraizados em contextos de desigualdade racial e socioeconômica.[226] Ao utilizar dados que refletem discriminações sistêmicas, o algoritmo acaba perpetuando essas injustiças, afetando negativamente a vida de indivíduos e comunidades.[227]

O caso do COMPAS destaca os perigos de confiar cegamente em sistemas algorítmicos sem transparência e sem considerar os vieses inerentes aos dados

---

[222] ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica, 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Acesso em: 10 out. 2023.
[223] ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica, 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Acesso em: 10 out. 2023.
[224] ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica, 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Acesso em: 10 out. 2023.
[225] ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica, 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Acesso em: 10 out. 2023.
[226] FLORES, A. W.; BECHTEL, K.; LOWENKAMP, C. False Positives, False Negatives, and False Analyses. Federal Probation, v. 80, n. 2, p. 38-46, 2016.
[227] FLORES, A. W.; BECHTEL, K.; LOWENKAMP, C. False Positives, False Negatives, and False Analyses. Federal Probation, v. 80, n. 2, p. 38-46, 2016.

de entrada.[228] Defensores dos direitos civis e pesquisadores têm exigido maior escrutínio e regulamentação sobre o uso de IA no sistema judicial, enfatizando a necessidade de desenvolver algoritmos justos e responsáveis.[229]

7.3. Sistemas de Reconhecimento Facial e Vieses Raciais e de Gênero

Estudos demonstraram que sistemas de reconhecimento facial apresentam taxas de erro significativamente maiores ao identificar mulheres e pessoas negras.[230] Esses sistemas são frequentemente treinados com conjuntos de dados desbalanceados, nos quais homens brancos estão super-representados. Como resultado, a precisão do reconhecimento diminui para grupos sub-representados, levando a potenciais erros e injustiças, especialmente quando utilizados por forças de segurança.[231]

Sistemas de reconhecimento facial têm sido amplamente adotados em diversas aplicações, desde desbloqueio de smartphones até vigilância por órgãos de segurança.[232] No entanto, pesquisas indicam que esses sistemas apresentam taxas de erro desproporcionais quando se trata de identificar mulheres e pessoas negras, em comparação com homens brancos.[233]

---

[228] KIRCHNER, L. When Discrimination Is Baked Into Algorithms. The Atlantic, 2016. Disponível em: https://www.theatlantic.com/business/archive/2016/09/discrimination-algorithms-disparate-impact/501329/ . Acesso em: 13 out. 2024.
[229] KIRCHNER, L. When Discrimination Is Baked Into Algorithms. The Atlantic, 2016. Disponível em: https://www.theatlantic.com/business/archive/2016/09/discrimination-algorithms-disparate-impact/501329/ . Acesso em: 13 out. 2024.
[230] BUOLAMWINI, J.; GEBRU, T. Gender shades: intersectional accuracy disparities in commercial gender classification. In: Conference on Fairness, Accountability and Transparency, 2018. Proceedings p.77-91.
[231] BUOLAMWINI, J.; GEBRU, T. Gender shades: intersectional accuracy disparities in commercial gender classification. In: Conference on Fairness, Accountability and Transparency, 2018. Proceedings p.77-91.
[232] GARVIE, C.; BEDOYA, A. M.; FRANKLE, J. The Perpetual Line-Up: Unregulated Police Face Recognition in America. Georgetown Law, Center on Privacy & Technology, 2016.
[233] GARVIE, C.; BEDOYA, A. M.; FRANKLE, J. The Perpetual Line-Up: Unregulated Police Face Recognition in America. Georgetown Law, Center on Privacy & Technology, 2016.

Um estudo relevou que três sistemas comerciais de reconhecimento facial apresentavam taxas de erro quase nulas ao identificar homens de pele clara, mas essas taxas aumentavam significativamente para mulheres de pele escura, chegando a até 34,7% de erro.[234] Essa disparidade é atribuída ao desequilíbrio nos conjuntos de dados utilizados para treinar os algoritmos, nos quais imagens de homens brancos estão super-representadas.[235]

As consequências desses vieses são preocupantes, especialmente quando esses sistemas são utilizados por forças de segurança e órgãos governamentais.[236] Erros na identificação podem levar a detenções injustas, violações de privacidade e reforço de práticas discriminatórias, afetando desproporcionalmente comunidades marginalizadas.[237]

Para mitigar esses problemas, pesquisadores e organizações têm defendido o desenvolvimento de conjuntos de dados mais diversos e representativos, além da implementação de regulamentações que garantam a transparência e a responsabilidade no uso de tecnologias de reconhecimento facial.[238] Empresas como IBM, Microsoft e Amazon anunciaram medidas para melhorar a precisão de seus sistemas e suspender ou limitar a venda dessas tecnologias para aplicações policiais.[239]

---

[234] BUOLAMWINI, J.; GEBRU, T. Gender shades: intersectional accuracy disparities in commercial gender classification. In: Conference on Fairness, Accountability and Transparency, 2018. Proceedings p.77-91.
[235] BUOLAMWINI, J.; GEBRU, T. Gender shades: intersectional accuracy disparities in commercial gender classification. In: Conference on Fairness, Accountability and Transparency, 2018. Proceedings p.77-91.
[236] LESEMAN, S.; SEIDMAN, D. Facial Recognition Technology: Potential Risks and Policy Considerations. Congressional Research Service, 2019.
[237] LESEMAN, S.; SEIDMAN, D. Facial Recognition Technology: Potential Risks and Policy Considerations. Congressional Research Service, 2019.
[238] HARDY, S. et al. Mitigating Bias in AI-Enabled Face Recognition. IBM Journal of Research and Development, v. 63, n. 4/5, p. 3:1-3:9, 2019.
[239] HARDY, S. et al. Mitigating Bias in AI-Enabled Face Recognition. IBM Journal of Research and Development, v. 63, n. 4/5, p. 3:1-3:9, 2019.

7.4. Google Translate e o Reforço de Estereótipos

O Google Translate foi criticado por reforçar estereótipos de gênero em traduções automáticas.[240] Por exemplo, ao traduzir frases de línguas sem gênero gramatical definido para o inglês, o sistema tendia a associar profissões tradicionalmente masculinas ao pronome "he" (ele) e profissões tradicionalmente femininas ao pronome "she" (ela). Isso ocorre devido aos vieses presentes nos dados de treinamento, que refletem desequilíbrios de gênero no mercado de trabalho.[241]

O Google Translate, amplamente utilizado para tradução automática, tem sido criticado por reforçar estereótipos de gênero em suas traduções.[242] Quando traduz frases de idiomas sem gênero gramatical definido para idiomas que exigem a especificação de gênero, o sistema tende a atribuir profissões e características de acordo com estereótipos tradicionais.[243]

Por exemplo, ao traduzir do turco para o inglês, frases como "O (a) médico(a) está aqui" são traduzidas como "He is here", enquanto "A (o) enfermeira(o) está aqui" torna-se "She is here".[244] Isso ocorre porque o algoritmo aprende a partir de grandes quantidades de texto disponível na internet, onde tais estereótipos estão presentes e são reproduzidos.[245]

---

[240] PRATES, M. O. R.; AVILA, S.; LAMB, L. C. Assessing gender bias in machine translation: A case study with Google Translate. Neural Computing and Applications, v. 32, n. 10, p. 6363-6381, 2020.
[241] PRATES, M. O. R.; AVILA, S.; LAMB, L. C. Assessing gender bias in machine translation: A case study with Google Translate. Neural Computing and Applications, v. 32, n. 10, p. 6363-6381, 2020.
[242] PRATES, M. O. R.; AVILA, S.; LAMB, L. C. Assessing gender bias in machine translation: A case study with Google Translate. Neural Computing and Applications, v. 32, n. 10, p. 6363-6381, 2020.
[243] PRATES, M. O. R.; AVILA, S.; LAMB, L. C. Assessing gender bias in machine translation: A case study with Google Translate. Neural Computing and Applications, v. 32, n. 10, p. 6363-6381, 2020.
[244] PRATES, M. O. R.; AVILA, S.; LAMB, L. C. Assessing gender bias in machine translation: A case study with Google Translate. Neural Computing and Applications, v. 32, n. 10, p. 6363-6381, 2020.
[245] PRATES, M. O. R.; AVILA, S.; LAMB, L. C. Assessing gender bias in machine translation: A case study with Google Translate. Neural Computing and Applications, v. 32, n. 10, p. 6363-6381, 2020.

Essa reprodução de vieses pode influenciar a percepção pública sobre papéis de gênero, reforçando ideias preconcebidas sobre quais profissões são adequadas para homens ou mulheres.[246] Além disso, pode impactar negativamente a representação de gênero em materiais educativos e profissionais que dependem de traduções automáticas.[247]

Para mitigar esses efeitos, o Google tem implementado esforços para oferecer traduções neutras ou múltiplas opções de gênero quando apropriado.[248] No entanto, o desafio permanece em equilibrar a precisão linguística com a responsabilidade social de não perpetuar estereótipos prejudiciais.[249]

7.5. Algoritmos em Plataformas de Mídia Social

Algoritmos de recomendação em plataformas como Facebook e YouTube têm sido acusados de criar "bolhas de filtro" que reforçam preconceitos e polarizações sociais.[250] Além disso, conteúdos misóginos e racistas podem ser amplificados por esses algoritmos, alcançando um público maior e normalizando discursos de ódio.[251]

Plataformas de mídia social utilizam algoritmos de recomendação para personalizar o conteúdo apresentado aos usuários, visando aumentar o

---

[246] CALISKAN, A.; BRYSON, J. J.; NARAYANAN, A. Semantics derived automatically from language corpora contain human-like biases. Science, v. 356, n. 6334, p. 183-186, 2017.
[247] CALISKAN, A.; BRYSON, J. J.; NARAYANAN, A. Semantics derived automatically from language corpora contain human-like biases. Science, v. 356, n. 6334, p. 183-186, 2017.
[248] GOOGLE. Reducing gender bias in Google Translate. Google AI Blog, 2018. Disponível em: https://ai.googleblog.com/2018/12/reducing-gender-bias-in-google.html . Acesso em: 10 out. 2023.
[249] GOOGLE. Reducing gender bias in Google Translate. Google AI Blog, 2018. Disponível em: https://ai.googleblog.com/2018/12/reducing-gender-bias-in-google.html . Acesso em: 10 out. 2023.
[250] PARISER, E. The Filter Bubble: What the Internet is Hiding from You. New York: Penguin Press, 2011.
[251] GILLESPIE, T. Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media. New Haven: Yale University Press, 2018.

engajamento.[252] No entanto, esses algoritmos podem criar "bolhas de filtro", nas quais os usuários são expostos apenas a informações e opiniões que reforçam suas crenças existentes, levando à polarização social.[253]

Além disso, conteúdos misóginos, racistas e extremistas podem ser amplificados por esses algoritmos, alcançando um público maior e normalizando discursos de ódio.[254] A lógica de maximizar o tempo de visualização e a interação muitas vezes favorece conteúdo sensacionalista ou controverso, independentemente de seu impacto social negativo.[255]

Casos de desinformação e manipulação de opiniões através de bots e campanhas coordenadas também evidenciam os riscos associados a esses sistemas.[256] Grupos extremistas têm explorado essas plataformas para recrutar membros e disseminar ideologias discriminatórias.[257]

Para enfrentar esses desafios, há um crescente debate sobre a responsabilidade das empresas de tecnologia em moderar o conteúdo e ajustar seus algoritmos para evitar a propagação de conteúdo prejudicial.[258] Medidas

---

[252] PARISER, E. The Filter Bubble: What the Internet is Hiding from You. New York: Penguin Press, 2011.
[253] PARISER, E. The Filter Bubble: What the Internet is Hiding from You. New York: Penguin Press, 2011.
[254] GILLESPIE, T. Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media. New Haven: Yale University Press, 2018.
[255] GILLESPIE, T. Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media. New Haven: Yale University Press, 2018.
[256] FERRARA, E. et al. The Rise of Social Bots. Communications of the ACM, v. 59, n. 7, p. 96-104, 2016.
[257] FERRARA, E. et al. The Rise of Social Bots. Communications of the ACM, v. 59, n. 7, p. 96-104, 2016.
[258] ZUCKERMAN, E. Q&A: Ethan Zuckerman on the Failure of Facebook's 'Good Intentions'. MIT News, 2019. Disponível em: http://news.mit.edu/2019/qa-ethan-zuckerman-facebook-good-intentions-0312 . Acesso em: 13 out. 2024.

como transparência algorítmica, regulamentação governamental e iniciativas de alfabetização midiática são propostas para mitigar os efeitos negativos.[259]

7.6. Sistemas de Avaliação de Crédito e Discriminação Financeira

Algoritmos utilizados por instituições financeiras para avaliação de crédito podem discriminar indiretamente mulheres e minorias étnicas.[260] Fatores aparentemente neutros, como histórico de endereço ou padrões de consumo, podem servir como proxies para características protegidas, levando à exclusão de grupos marginalizados do acesso a serviços financeiros.[261]

Sistemas de avaliação de crédito utilizam algoritmos para analisar dados financeiros e determinar a elegibilidade de indivíduos para empréstimos e outros serviços.[262] No entanto, esses algoritmos podem discriminar indiretamente mulheres e minorias étnicas.[263] Fatores aparentemente neutros, como histórico de endereço ou padrões de consumo, podem servir como proxies para características protegidas, levando à exclusão de grupos marginalizados do acesso a serviços financeiros.[264]

Em 2019, o Apple Card, um cartão de crédito lançado pela Apple em parceria com o Goldman Sachs, foi acusado de discriminação de gênero em seu algoritmo de determinação de limite de crédito.[265] Usuários relataram que

---

[259] ZUCKERMAN, E. Q&A: Ethan Zuckerman on the Failure of Facebook's 'Good Intentions'. MIT News, 2019. Disponível em: http://news.mit.edu/2019/qa-ethan-zuckerman-facebook-good-intentions-0312 . Acesso em: 13 out. 2024.
[260] HURLEY, M.; ADEBAYO, J. Credit scoring in the era of big data. Yale Journal of Law & Technology, v. 18, p. 148-216, 2016.
[261] HURLEY, M.; ADEBAYO, J. Credit scoring in the era of big data. Yale Journal of Law & Technology, v. 18, p. 148-216, 2016.
[262] HURLEY, M.; ADEBAYO, J. Credit scoring in the era of big data. Yale Journal of Law & Technology, v. 18, p. 148-216, 2016.
[263] HURLEY, M.; ADEBAYO, J. Credit scoring in the era of big data. Yale Journal of Law & Technology, v. 18, p. 148-216, 2016.
[264] HURLEY, M.; ADEBAYO, J. Credit scoring in the era of big data. Yale Journal of Law & Technology, v. 18, p. 148-216, 2016.
[265] BBC NEWS. Apple Card investigated after claims of gender bias. BBC News, 2019. Disponível em: https://www.bbc.com/news/business-50365609 . Acesso em: 13 out. 2024.

mulheres estavam recebendo limites de crédito significativamente menores do que homens com perfis financeiros semelhantes.[266]

O desenvolvedor David Heinemeier Hansson denunciou publicamente que recebeu um limite de crédito 20 vezes maior que o de sua esposa, apesar de ela ter uma melhor pontuação de crédito.[267] O cofundador da Apple, Steve Wozniak, também relatou experiência similar, indicando que o problema não era um caso isolado.[268]

Essas alegações levantaram preocupações sobre a opacidade dos algoritmos utilizados para determinar limites de crédito e como eles podem incorporar vieses discriminatórios.[269] A autoridade reguladora do Departamento de Serviços Financeiros do Estado de Nova York iniciou uma investigação sobre o caso, enfatizando a necessidade de transparência e equidade nos sistemas automatizados de crédito.[270]

O Goldman Sachs respondeu afirmando que não considera gênero em suas decisões de crédito e que analisaria os casos reportados.[271] No entanto, o

---

[266] BBC NEWS. Apple Card investigated after claims of gender bias. BBC News, 2019. Disponível em: https://www.bbc.com/news/business-50365609 . Acesso em: 13 out. 2024.
[267] VIGDOR, Neil. Apple Card Investigated After Gender Discrimination Complaints. The New York Times, 10 nov. 2019. Disponível em: https://www.nytimes.com/2019/11/10/business/Apple-credit-card-investigation.html . Acesso em: 13 out. 2024; DAY, M. Apple Card algorithm sparks gender bias allegations. Financial Times, 2019. Disponível em: https://www.ft.com/content/3d4b2d8e-05d3-11ea-a984-fbbacad9e7dd . Acesso em: 13 out. 2024.
[268] VIGDOR, Neil. Apple Card Investigated After Gender Discrimination Complaints. The New York Times, 10 nov. 2019. Disponível em: https://www.nytimes.com/2019/11/10/business/Apple-credit-card-investigation.html . Acesso em: 13 out. 2024; DAY, M. Apple Card algorithm sparks gender bias allegations. Financial Times, 2019. Disponível em: https://www.ft.com/content/3d4b2d8e-05d3-11ea-a984-fbbacad9e7dd . Acesso em: 13 out. 2024.
[269] DAY, M. Apple Card algorithm sparks gender bias allegations. Financial Times, 2019. Disponível em: https://www.ft.com/content/3d4b2d8e-05d3-11ea-a984-fbbacad9e7dd . Acesso em: 13 out. 2024.
[270] JAMES, L. DFS to Investigate Apple Card for Gender Discrimination. New York State Department of Financial Services, 2019. Disponível em: https://www.dfs.ny.gov/reports_and_publications/press_releases/pr1911101 . Acesso em: 13 out. 2024.
[271] JAMES, L. DFS to Investigate Apple Card for Gender Discrimination. New York State Department of Financial Services, 2019. Disponível em:

incidente destacou como algoritmos financeiros podem perpetuar desigualdades de gênero, mesmo sem intenção explícita, ao basear-se em dados históricos que refletem vieses sistêmicos.[272]

Um estudo conduzido pela Universidade da Califórnia em Berkeley revelou que algoritmos utilizados para determinar taxas de juros em empréstimos imobiliários cobravam taxas mais altas de minorias étnicas em comparação com candidatos brancos com perfis financeiros semelhantes.[273] A pesquisa analisou mais de 7 milhões de aplicações de empréstimo e descobriu que, em média, latinos e afro-americanos pagavam até 0,079% a mais em taxas de juros.[274]

Essa diferença aparentemente pequena resulta em milhares de dólares a mais pagos ao longo do prazo do empréstimo, impactando significativamente a capacidade financeira dessas comunidades.[275] O estudo sugeriu que os algoritmos estavam considerando fatores correlacionados com raça, mesmo que a raça em si não fosse um parâmetro explícito.[276]

Esse caso evidencia como sistemas automatizados podem perpetuar discriminações históricas presentes no mercado imobiliário, onde minorias étnicas tradicionalmente enfrentam barreiras para aquisição de propriedades.[277]

---

https://www.dfs.ny.gov/reports_and_publications/press_releases/pr1911101 . Acesso em: 13 out. 2024.

[272] SIDDHARTHA, V. Apple Card's Gender Bias: What Really Happened. Forbes, 2019. Disponível em: https://www.forbes.com/sites/siddharthavarma/2019/11/11/apple-cards-gender-bias-what-really-happened/ . Acesso em: 13 out. 2024.

[273] BARTLETT, R.; MORRISON, S.; NELSON, L.; RISKER, M. Consumer Lending Discrimination in the FinTech Era. NBER Working Paper No. 25943, 2019.

[274] BARTLETT, R.; MORRISON, S.; NELSON, L.; RISKER, M. Consumer Lending Discrimination in the FinTech Era. NBER Working Paper No. 25943, 2019.

[275] BARTLETT, R.; MORRISON, S.; NELSON, L.; RISKER, M. Consumer Lending Discrimination in the FinTech Era. NBER Working Paper No. 25943, 2019.

[276] BARTLETT, R.; MORRISON, S.; NELSON, L.; RISKER, M. Consumer Lending Discrimination in the FinTech Era. NBER Working Paper No. 25943, 2019.

[277] BHARGAVA, V.; LIVINGSTON, J. Regulating Algorithms in Consumer Credit. Harvard Journal of Law & Technology, v. 33, n. 1, p. 167-214, 2019.

A falta de transparência nos critérios utilizados pelos algoritmos dificulta a identificação e correção desses vieses.[278]

7.7. Chatbots e Discurso Discriminatório

Em 2016, a Microsoft lançou o chatbot Tay no Twitter, que rapidamente começou a reproduzir discursos racistas e misóginos após interagir com usuários.[279] Isso ocorreu porque o chatbot aprendeu a partir das mensagens enviadas pelos usuários, sem filtros adequados para conteúdos inadequados.[280]

O chatbot Tay foi projetado para aprender interativamente a partir das interações com os usuários.[281] O objetivo era demonstrar as capacidades de aprendizado de máquina em linguagem natural e interação social. No entanto, em poucas horas, Tay começou a reproduzir e disseminar discursos racistas, misóginos e de ódio, resultado da manipulação por parte de usuários que enviaram mensagens com conteúdo inadequado.[282]

A falta de filtros e salvaguardas permitiu que o chatbot assimilasse e replicasse esses padrões de linguagem. Este incidente expôs os desafios éticos e técnicos na implementação de sistemas de IA que aprendem diretamente de interações públicas.[283] Sem mecanismos para detectar e evitar conteúdo

---

[278] BHARGAVA, V.; LIVINGSTON, J. Regulating Algorithms in Consumer Credit. Harvard Journal of Law & Technology, v. 33, n. 1, p. 167-214, 2019.
[279] NEFF, G.; NAGY, P. Talking to bots: Symbiotic agency and the case of Tay. International Journal of Communication, v. 10, p. 4915-4931, 2016.
[280] NEFF, G.; NAGY, P. Talking to bots: Symbiotic agency and the case of Tay. International Journal of Communication, v. 10, p. 4915-4931, 2016.
[281] NEFF, G.; NAGY, P. Talking to bots: Symbiotic agency and the case of Tay. International Journal of Communication, v. 10, p. 4915-4931, 2016.
[282] NEFF, G.; NAGY, P. Talking to bots: Symbiotic agency and the case of Tay. International Journal of Communication, v. 10, p. 4915-4931, 2016.
[283] PARRY, D. et al. Understanding the Limitations of AI: When Tay Tweets Go Bad. In: ACM CHI Conference on Human Factors in Computing Systems, 2016.

prejudicial, tais sistemas podem amplificar discursos discriminatórios e causar danos à reputação das empresas e à sociedade em geral.[284]

Como resposta, a Microsoft retirou Tay do ar e posteriormente lançou o chatbot Zo, com melhorias para evitar a repetição dos erros anteriores.[285] O caso de Tay destaca a importância de incorporar princípios de ética e responsabilidade no design de IA, incluindo a moderação de conteúdo e o treinamento para evitar a reprodução de vieses e discursos de ódio.[286]

8. Abordagens Críticas: Materialismo Histórico-Dialético, Escola de Frankfurt, Feminismo e Sociologia da Ciência e Tecnologia

A inclusão de diferentes perspectivas teóricas na análise do viés de discriminação algorítmica é fundamental para enriquecer a compreensão das raízes e das consequências desse fenômeno. Cada abordagem teórica oferece um conjunto distinto de ferramentas conceituais para explorar como os sistemas algorítmicos refletem e reforçam as desigualdades sociais. Ao considerar diferentes perspectivas, como o materialismo histórico-dialético, o feminismo e a sociologia da ciência e da tecnologia, torna-se possível revelar as múltiplas camadas de poder, opressão e exclusão incorporadas nas tecnologias contemporâneas.

Essa diversidade de enfoques permite que a análise vá além de uma visão técnica ou isolada do problema, contemplando os contextos econômicos, políticos e sociais nos quais os algoritmos são desenvolvidos e aplicados. Considerar essas diferentes lentes teóricas é essencial para identificar não

---

[284] PARRY, D. et al. Understanding the Limitations of AI: When Tay Tweets Go Bad. In: ACM CHI Conference on Human Factors in Computing Systems, 2016.
[285] VINCENT, J. Microsoft's racist chatbot returns with new rules. The Verge, 2017. Disponível em: https://www.theverge.com/2017/8/8/16113530/microsoft-zo-ai-chatbot-returns . Acesso em: 13 out. 2024.
[286] VINCENT, J. Microsoft's racist chatbot returns with new rules. The Verge, 2017. Disponível em: https://www.theverge.com/2017/8/8/16113530/microsoft-zo-ai-chatbot-returns . Acesso em: 13 out. 2024.

apenas os sintomas, mas também as causas estruturais e históricas da discriminação algorítmica, propondo soluções que levem em conta a complexidade do tema.

8.1. Materialismo Histórico-Dialético

O materialismo histórico-dialético é uma abordagem teórica desenvolvida por Karl Marx e Friedrich Engels que busca compreender a evolução das sociedades humanas a partir das condições materiais de existência e das contradições inerentes às relações econômicas.[287] Fundamentado no materialismo filosófico, ele propõe que a realidade material, especialmente as relações de produção e as forças produtivas, determina a estrutura social, política e ideológica de uma sociedade.[288]

Sob essa perspectiva, a história é vista como um processo dialético e dinâmico, onde as mudanças sociais ocorrem por meio de conflitos entre classes sociais com interesses antagônicos.[289] A dialética, inspirada na filosofia hegeliana, é utilizada para explicar que o desenvolvimento social não é linear, mas resulta da tensão entre forças opostas que geram transformações qualitativas.[290] Essas contradições são motoras da história, levando a avanços ou retrocessos nas estruturas sociais.[291]

No centro dessa teoria está a ideia de que a infraestrutura econômica (a base material da sociedade) determina a superestrutura (as instituições políticas, jurídicas, ideológicas e culturais).[292] Isso significa que as formas de organização social, as leis, a cultura e até mesmo a consciência coletiva são

[287] MARX, K.; ENGELS, F. A Ideologia Alemã. São Paulo: Boitempo, 2007.
[288] MARX, Karl. O Capital: Crítica da Economia Política. São Paulo: Nova Cultural, 1988.
[289] MARX, K.; ENGELS, F. A Ideologia Alemã. São Paulo: Boitempo, 2007.
[290] HEGEL, G. W. F. Fenomenologia do Espírito. Petrópolis: Vozes, 2012.
[291] HEGEL, G. W. F. Fenomenologia do Espírito. Petrópolis: Vozes, 2012.
[292] MARX, K. O Capital: Crítica da Economia Política. São Paulo: Nova Cultural, 1988.

moldadas pelas relações econômicas predominantes.[293] A ideologia dominante em uma sociedade reflete os interesses da classe que detém o controle dos meios de produção.[294]

Um conceito fundamental relacionado é o da alienação, que ocorre quando os trabalhadores são separados dos produtos de seu trabalho, do processo produtivo, de sua própria essência humana e de outros seres humanos.[295] No sistema capitalista, o trabalho é transformado em mercadoria, e o trabalhador perde o controle sobre sua atividade e criatividade, resultando em uma desconexão entre o indivíduo e sua natureza humana.[296] Isso gera um sentimento de despersonalização e perda de significado no trabalho.[297]

Outro conceito relevante é o de reificação, apresentado por Georg Lukács, que se refere à transformação de relações sociais em coisas, onde processos sociais e humanos são percebidos como objetos ou commodities.[298] No contexto capitalista, as relações humanas são mediadas por mercadorias e valores de troca, obscurecendo a natureza social e humana dessas interações.[299]

Aplicando esses conceitos ao contexto contemporâneo da inteligência artificial e das Big Techs, observa-se que o desenvolvimento tecnológico está profundamente inserido nas relações de produção capitalistas.[300] As grandes corporações tecnológicas controlam os meios de produção digital—infraestrutura, plataformas e dados—e direcionam o desenvolvimento da IA

---

[293] MARX, K. O Capital: Crítica da Economia Política. São Paulo: Nova Cultural, 1988.
[294] MARX, K.; ENGELS, F. A Ideologia Alemã. São Paulo: Boitempo, 2007.
[295] MARX, K. Manuscritos Econômico-Filosóficos. São Paulo: Boitempo, 2010.
[296] MARX, K. Manuscritos Econômico-Filosóficos. São Paulo: Boitempo, 2010.
[297] MARX, K. Manuscritos Econômico-Filosóficos. São Paulo: Boitempo, 2010.
[298] LUKÁCS, G. História e Consciência de Classe. São Paulo: Martins Fontes, 2012.
[299] LUKÁCS, G. História e Consciência de Classe. São Paulo: Martins Fontes, 2012.
[300] ZUBOFF, S. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs, 2019.

conforme seus interesses econômicos, visando maximizar lucros e consolidar seu domínio de mercado.[301]

Os algoritmos de IA são treinados com dados que refletem a realidade social existente, incluindo desigualdades e preconceitos históricos.[302] Isso resulta em vieses algorítmicos que podem perpetuar ou até amplificar discriminações de gênero, raça e classe.[303] Sob a perspectiva do materialismo histórico-dialético, esses vieses não são meras falhas técnicas, mas manifestações das contradições do sistema capitalista, onde as relações de produção influenciam o desenvolvimento tecnológico de forma a servir aos interesses da classe dominante.[304]

A alienação é evidente quando consideramos que tanto os trabalhadores envolvidos no desenvolvimento da IA quanto os usuários finais têm pouco controle sobre como a tecnologia é utilizada ou sobre os impactos sociais que ela gera.[305] Os dados pessoais dos usuários são coletados e monetizados sem consentimento pleno, transformando indivíduos em mercadorias dentro do capitalismo de vigilância.[306] Isso reforça a alienação, pois as pessoas se tornam objetos dentro do processo produtivo tecnológico.

A reificação ocorre quando as interações humanas são mediadas por algoritmos que tratam indivíduos como conjuntos de dados a serem analisados

---

[301] ZUBOFF, S. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs, 2019.
[302] NOBLE, S. U. Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press, 2018.
[303] NOBLE, S. U. Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press, 2018.
[304] MARX, K. O Capital: Crítica da Economia Política. São Paulo: Nova Cultural, 1988.
[305] ZUBOFF, S. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs, 2019.
[306] ZUBOFF, S. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs, 2019.

e manipulados.[307] As decisões tomadas por sistemas de IA podem desconsiderar a complexidade humana, reduzindo pessoas a perfis estatísticos. Isso pode ser observado em práticas como a segmentação de mercado, a personalização de conteúdos e a avaliação de crédito automatizada.

Além disso, a concentração de poder nas mãos das Big Techs permite que elas influenciem a superestrutura ideológica, moldando cultura, política e normas sociais por meio do controle de plataformas de comunicação e fluxos de informação.[308] Isso reflete a ideia marxista de que a classe dominante também domina a produção ideológica, utilizando sua posição para perpetuar relações de produção que lhe são favoráveis.[309]

A lógica de maximização de capital dessas empresas leva à priorização de mercados lucrativos, muitas vezes negligenciando necessidades sociais ou perpetuando desigualdades.[310] Por exemplo, o desenvolvimento de tecnologias que não consideram a diversidade cultural e linguística marginaliza comunidades que não representam retornos financeiros imediatos, perpetuando uma forma de colonialismo digital.[311]

8.2. A Escola de Frankfurt e a Crítica da Razão Instrumental

A Escola de Frankfurt, representada por pensadores como Max Horkheimer, Theodor Adorno e Herbert Marcuse, oferece uma análise crítica da sociedade contemporânea, especialmente em relação ao papel da indústria cultural e da tecnologia na manutenção das estruturas de dominação.[312] A

[307] LUKÁCS, G. História e Consciência de Classe. São Paulo: Martins Fontes, 2012.
[308] ZUBOFF, S. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs, 2019.
[309] MARX, K.; ENGELS, F. A Ideologia Alemã. São Paulo: Boitempo, 2007.
[310] NOBLE, S. U. Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press, 2018.
[311] COULDRY, N.; MEJIAS, U. A. The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism. Stanford: Stanford University Press, 2019.
[312] HORKHEIMER, M.; ADORNO, T. W. Dialética do Esclarecimento. Rio de Janeiro: Zahar, 1985.

crítica à razão instrumental é central nessa abordagem, apontando como a racionalidade técnica é utilizada para controlar e manipular tanto a natureza quanto os seres humanos, subordinando-os aos interesses do capital.[313]

Segundo Horkheimer e Adorno, a razão instrumental transforma a racionalidade em um meio para atingir fins econômicos e administrativos, desconsiderando valores éticos e emancipatórios.[314] No contexto das Big Techs e da IA, essa racionalidade instrumental se manifesta na utilização de algoritmos para otimizar lucros e eficiência, frequentemente em detrimento de considerações sociais e humanas. A tecnologia deixa de ser um meio de libertação e passa a servir como instrumento de controle e conformismo.

Herbert Marcuse argumenta que a tecnologia, sob o capitalismo avançado, contribui para a criação de uma sociedade unidimensional, onde a capacidade crítica dos indivíduos é suprimida.[315] A indústria tecnológica produz necessidades artificiais e promove uma falsa consciência, levando as pessoas a aceitar passivamente o status quo. A IA, nesse sentido, pode ser vista como um mecanismo que reforça padrões de consumo e comportamento, limitando a autonomia individual e a diversidade de pensamento.

A análise da Escola de Frankfurt também ressalta o papel da indústria cultural na disseminação de ideologias que legitimam as relações de poder existentes.[316] As Big Techs, ao controlarem plataformas de mídia social e distribuição de conteúdo, influenciam a formação de opiniões e valores culturais. Isso possibilita a manipulação de informações e a perpetuação de

---

[313] HORKHEIMER, M.; ADORNO, T. W. Dialética do Esclarecimento. Rio de Janeiro: Zahar, 1985.
[314] HORKHEIMER, M.; ADORNO, T. W. Dialética do Esclarecimento. Rio de Janeiro: Zahar, 1985.
[315] MARCUSE, H. One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society. Boston: Beacon Press, 1964.
[316] HORKHEIMER, M.; ADORNO, T. W. Dialética do Esclarecimento. Rio de Janeiro: Zahar, 1985.

estereótipos, contribuindo para a manutenção das estruturas de dominação e a redução do espaço para a crítica e a resistência.

8.3. Pós-Estruturalismo, Poder e Algoritmos

O pós-estruturalismo, particularmente nas obras de Michel Foucault, oferece um conjunto de conceitos fundamentais para entender a relação entre poder, saber e tecnologia. Foucault argumenta que o poder não é apenas uma força repressiva exercida por instituições, mas algo que circula por toda a sociedade, operando de maneira capilar e incorporado em práticas cotidianas, relações sociais e discursos.[317] Em vez de ser centralizado em uma autoridade, o poder é distribuído através de múltiplos pontos e mecanismos, moldando comportamentos e conhecimentos em todos os níveis da vida social.

Um dos conceitos centrais de Foucault é o de **poder-saber**, que explora a relação entre a produção de conhecimento e o exercício do poder. Para Foucault, o saber é uma forma de poder, e as formas de conhecimento que são legitimadas por uma sociedade são inseparáveis das estruturas de poder que as sustentam.[318] No contexto da inteligência artificial, isso significa que os algoritmos, ao processarem e analisarem dados para gerar conhecimento, também estão exercendo poder ao definir quais informações são valiosas e quais decisões devem ser tomadas com base nesses dados. Por exemplo, os algoritmos que categorizam e monitoram usuários de redes sociais ou que avaliam a concessão de crédito reproduzem formas de controle sobre indivíduos, refletindo e reforçando as hierarquias e desigualdades preexistentes.

[317] FOUCAULT, M. Nascimento da Biopolítica: Curso no Collège de France (1978-1979). São Paulo: Martins Fontes, 2007.
[318] FOUCAULT, M. A Arqueologia do Saber. Rio de Janeiro: Forense Universitária, 2012.

O conceito de **biopolítica**, desenvolvido por Foucault, refere-se ao exercício do poder sobre a vida e a gestão de populações. A biopolítica emerge como uma forma de poder que regula a vida em um nível coletivo, intervindo na saúde, na reprodução, na demografia e na segurança das populações.[319] A IA pode ser vista como uma extensão dessa biopolítica, pois sistemas algorítmicos são usados para monitorar, prever e controlar comportamentos populacionais em grande escala, muitas vezes reforçando desigualdades de gênero, raça e classe. Um exemplo disso são os sistemas de IA utilizados para prever crimes, que frequentemente reproduzem preconceitos raciais, submetendo determinadas populações a maior vigilância e controle.

Governamentalidade é outro conceito central em Foucault, que diz respeito às formas de governar as condutas dos indivíduos e das populações através de estratégias racionais e técnicas de controle.[320] Na era contemporânea, a governamentalidade está intimamente ligada à lógica neoliberal, onde a eficiência econômica e o mercado são os principais critérios de governança. As grandes corporações de tecnologia, como as Big Techs, exercem uma forma de governamentalidade ao utilizar algoritmos para regular o comportamento dos usuários em suas plataformas e para otimizar o controle de dados pessoais, alinhando o uso da tecnologia a interesses econômicos. A IA, nesse contexto, torna-se uma ferramenta de governança que molda o comportamento dos indivíduos de maneira invisível, mas poderosa, transformando dados pessoais em instrumentos de vigilância e controle.

---

[319] FOUCAULT, M. História da Sexualidade I: A Vontade de Saber. Rio de Janeiro: Graal, 1999.

[320] FOUCAULT, M. Nascimento da Biopolítica: Curso no Collège de France (1978-1979). São Paulo: Martins Fontes, 2007.

Os **dispositivos de segurança** são mecanismos que Foucault descreve como tecnologias de poder que buscam gerir e controlar riscos em sociedades modernas, regulando o comportamento das populações.[321] No contexto dos sistemas de IA, os algoritmos atuam como dispositivos de segurança ao serem empregados para prever comportamentos e gerir riscos sociais, como no caso de algoritmos de predição criminal ou avaliação de risco de crédito. No entanto, esses dispositivos frequentemente reproduzem discriminações e vulnerabilidades existentes, como no caso de algoritmos que discriminam minorias étnicas e raciais em julgamentos de reincidência criminal ou na concessão de empréstimos.

Por fim, o conceito de **disciplinaridade** complementa essa análise. Foucault argumenta que o poder disciplinar se manifesta por meio da vigilância, da normatização e da punição, moldando o comportamento dos indivíduos em diferentes instituições, como prisões, escolas e fábricas.[322] Na era da IA, o poder disciplinar se estende para o espaço digital, onde algoritmos monitoram e regulam as ações dos usuários, definindo padrões de comportamento desejáveis e marginalizando aqueles que se desviam desses padrões. Isso é visível, por exemplo, em plataformas digitais que utilizam IA para personalizar conteúdos, promover comportamentos de consumo ou excluir vozes dissidentes, reforçando uma conformidade social invisível, porém eficaz.

Dessa forma, o pós-estruturalismo, especialmente através das ideias de Foucault, oferece uma análise crítica do papel da IA e das Big Techs na contemporaneidade, mostrando como as tecnologias digitais reproduzem e

---

[321] FOUCAULT, M. Segurança, Território, População: Curso no Collège de France (1977-1978). São Paulo: Martins Fontes, 2008.

[322] FOUCAULT, M. Vigiar e Punir: Nascimento da Prisão. Petrópolis: Vozes, 1975.

amplificam mecanismos de poder e controle. A IA, vista sob essa perspectiva, não é apenas uma ferramenta técnica, mas um instrumento político que regula a vida social, disciplina os corpos e populações e reforça desigualdades preexistentes através de mecanismos invisíveis, mas profundos.

8.3. Feminismo e Teoria de Gênero: Desconstruindo Vieses e Discriminações na IA

As abordagens feministas e de teoria de gênero oferecem ferramentas fundamentais para analisar como as tecnologias, incluindo a IA, podem reproduzir e reforçar desigualdades de gênero.[323] O feminismo interseccional, proposto por Kimberlé Crenshaw, destaca como diferentes formas de opressão—gênero, raça, classe—se sobrepõem e intensificam a marginalização de grupos específicos.[324] Aplicando essa perspectiva, observa-se que os vieses algorítmicos na IA afetam desproporcionalmente mulheres e minorias étnicas.

A falta de diversidade nas equipes de desenvolvimento tecnológico contribui para a incorporação de vieses inconscientes nos sistemas de IA.[325] As experiências e perspectivas de mulheres e outros grupos sub-representados são frequentemente negligenciadas, resultando em tecnologias que não atendem às suas necessidades ou que os prejudicam diretamente. Isso reforça as estruturas patriarcais e a desigualdade de gênero presentes na sociedade, conforme descrito por Gilberto Freyre em relação ao patriarcado brasileiro.

Judith Butler argumenta que gênero é uma construção social performativa, e que as categorias fixas de gênero podem ser limitantes e

---

[323] HARAWAY, D. J. Simians, Cyborgs, and Women: The Reinvention of Nature. New York: Routledge, 1991.
[324] CRENSHAW, K. Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. Stanford Law Review, v. 43, n. 6, p. 1241-1299, 1991.
[325] CRIADO PEREZ, C. Invisible Women: Data Bias in a World Designed for Men. New York: Abrams Press, 2019.

excludentes.[326] Na IA, a utilização de categorias binárias e estereotipadas de gênero pode perpetuar preconceitos e ignorar a diversidade das identidades de gênero. Assistentes virtuais com vozes femininas e personalidades submissas, por exemplo, reforçam estereótipos de gênero e normalizam a objetificação das mulheres.

Além disso, a interseccionalidade ressalta que as discriminações não ocorrem de forma isolada.[327] Mulheres negras, por exemplo, enfrentam desafios específicos que combinam racismo e sexismo.[328] Sistemas de reconhecimento facial têm maior probabilidade de erros ao identificar mulheres negras, evidenciando como a IA pode reproduzir múltiplas camadas de discriminação.[329] Isso demonstra a necessidade de abordagens que considerem a complexidade das identidades e das experiências humanas no desenvolvimento tecnológico.

8.4. Sociologia da Ciência e Tecnologia

A Sociologia da Ciência e Tecnologia (STS) é um campo interdisciplinar que busca entender como fatores sociais, culturais e políticos influenciam a prática científica e o desenvolvimento tecnológico.[330] Diferente da visão tradicional que considera a ciência e a tecnologia como empreendimentos neutros e objetivos, a STS parte do princípio de que essas áreas são

[326] BUTLER, J. Gender Trouble: Feminism and the Subversion of Identity. New York: Routledge, 1990.
[327] CRENSHAW, K. Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. Stanford Law Review, v. 43, n. 6, p. 1241-1299, 1991.
[328] CRENSHAW, K. Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. Stanford Law Review, v. 43, n. 6, p. 1241-1299, 1991.
[329] BUOLAMWINI, J.; GEBRU, T. Gender shades: intersectional accuracy disparities in commercial gender classification. In: Conference on Fairness, Accountability and Transparency, 2018. Proceedings p.77-91.
[330] SISMONDO, S. An Introduction to Science and Technology Studies. Chichester: Wiley-Blackwell, 2010.

profundamente moldadas pelas estruturas sociais e valores da sociedade em que são produzidas.[331]

Os estudiosos da STS, como Bruno Latour, Steven Shapin e Thomas Kuhn, argumentam que o conhecimento científico e os avanços tecnológicos são resultados de processos sociais complexos, envolvendo negociações, redes de atores e construções culturais. Latour, por exemplo, através de sua teoria de *Actor-Network Theory (ANT),* enfatiza que tanto os seres humanos quanto os objetos técnicos (ou "não-humanos") são agentes no desenvolvimento científico e tecnológico.[332] Esse ponto de vista descentraliza a ação humana como única força motriz e considera que as tecnologias também influenciam as ações e decisões humanas, formando uma rede complexa de interações.

Outro ponto fundamental da STS é a ideia de que o progresso científico não é simplesmente uma acumulação de conhecimento, mas envolve rupturas, controvérsias e mudanças de paradigmas, como apontado por Thomas Kuhn em seu influente livro A Estrutura das Revoluções Científicas.[333] Kuhn sugere que o conhecimento científico avança através de crises e revoluções, onde um paradigma é substituído por outro, muitas vezes rompendo com as concepções anteriores da ciência e da realidade.[334]

A ciência, vista por esse prisma, não é um empreendimento puramente racional, mas um campo de disputas que envolve interesses econômicos, políticos e culturais. Steven Shapin e Simon Schaffer, em *Leviathan and the Air-Pump*, mostram como as descobertas científicas no século XVII foram

---

[331] SISMONDO, S. An Introduction to Science and Technology Studies. Chichester: Wiley-Blackwell, 2010.
[332] LATOUR, B. Reassembling the Social: An Introduction to Actor-Network-Theory. Oxford: Oxford University Press, 2005.
[333] KUHN, T. S. A Estrutura das Revoluções Científicas. São Paulo: Perspectiva, 1962.
[334] KUHN, T. S. A Estrutura das Revoluções Científicas. São Paulo: Perspectiva, 1962.

profundamente influenciadas pelas práticas políticas e sociais da época.[335] Eles argumentam que a ciência não pode ser separada das estruturas sociais, e que as instituições científicas, os laboratórios e as práticas experimentais estão imersos em contextos sociais que influenciam quais tipos de conhecimento são legitimados.[336]

No campo da tecnologia, a STS explora como as inovações tecnológicas são co-produzidas com valores sociais. Sheila Jasanoff, uma das principais teóricas da STS, desenvolveu o conceito de co-produção, argumentando que a ciência e a tecnologia não apenas refletem, mas também moldam os sistemas sociais e políticos.[337] A tecnologia, ao ser adotada e integrada à sociedade, transforma comportamentos, práticas e relações de poder, e, ao mesmo tempo, é moldada por esses elementos.

Aplicando a STS ao contexto da inteligência artificial (IA) e da discriminação algorítmica, podemos entender que os sistemas de IA não são apenas produtos de avanço técnico, mas também resultados de decisões sociais sobre quais dados utilizar, como treinar algoritmos e quais objetivos priorizar. Por exemplo, os vieses presentes nos algoritmos refletem as desigualdades sociais que já estão presentes nos dados e nos processos de coleta.[338] Assim, a STS nos fornece ferramentas para compreender que a discriminação algorítmica não é apenas uma questão técnica, mas também uma manifestação das estruturas sociais e das relações de poder existentes.

---

[335] SHAPIN, S.; SCHAFFER, S. Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life. Princeton: Princeton University Press, 1985.
[336] SHAPIN, S.; SCHAFFER, S. Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life. Princeton: Princeton University Press, 1985.
[337] JASANOFF, S. States of Knowledge: The Co-Production of Science and Social Order. London: Routledge, 2004.
[338] EUBANKS, V. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin's Press, 2018.

A STS também desafia a ideia de neutralidade científica ao examinar como as práticas tecnocientíficas podem servir aos interesses de determinados grupos sociais, frequentemente os mais poderosos.[339] No caso das grandes corporações tecnológicas (Big Techs), essas empresas utilizam a ciência de dados e a IA para fortalecer seus modelos de negócios, muitas vezes em detrimento de preocupações com igualdade e justiça social.[340] A STS, portanto, fornece uma lente crítica para analisar como essas tecnologias podem perpetuar desigualdades sociais e econômicas.

8.5. Construção Social da Tecnologia e a Inteligência Artificial

A teoria da Construção Social da Tecnologia (SCOT) é uma abordagem dentro da Sociologia da Ciência e Tecnologia que propõe que o desenvolvimento tecnológico não é um processo linear e determinista, mas é influenciado por fatores sociais, culturais e políticos.[341] Os principais teóricos dessa abordagem, Trevor Pinch e Wiebe Bijker, argumentam que a tecnologia é moldada por diferentes grupos sociais relevantes, que interagem e negociam significados e usos para os artefatos tecnológicos.[342]

Um conceito central na SCOT é o de flexibilidade interpretativa, que sugere que um artefato tecnológico pode ter diferentes significados e usos para diferentes grupos sociais.[343] Isso implica que as características técnicas não são

---

[339] LATOUR, B. Reassembling the Social: An Introduction to Actor-Network-Theory. Oxford: Oxford University Press, 2005.
[340] ZUBOFF, S. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs, 2019.
[341] PINCH, T. J.; BIJKER, W. E. The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other. Social Studies of Science, v. 14, n. 3, p. 399-441, 1984.
[342] PINCH, T. J.; BIJKER, W. E. The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other. Social Studies of Science, v. 14, n. 3, p. 399-441, 1984.
[343] PINCH, T. J.; BIJKER, W. E. The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other. Social Studies of Science, v. 14, n. 3, p. 399-441, 1984.

inerentes, mas resultam de processos de construção social.[344] Grupos sociais relevantes, como engenheiros, usuários, reguladores e empresas, contribuem para a definição de problemas e soluções tecnológicas, influenciando o design e a implementação de tecnologias.[345]

O processo de fechamento e estabilização ocorre quando um consenso é alcançado entre os grupos sociais sobre a forma e a função de uma tecnologia, reduzindo a flexibilidade interpretativa.[346] No entanto, esse fechamento pode marginalizar perspectivas de grupos menos poderosos, cujas necessidades e valores não são considerados no design final. Isso levanta questões sobre poder e inclusão no desenvolvimento tecnológico.

Langdon Winner contribui para o debate ao questionar se artefatos tecnológicos têm política.[347] Ele argumenta que tecnologias incorporam formas de poder e autoridade, podendo reforçar ou desafiar estruturas sociais existentes.[348] Essa perspectiva complementa a SCOT ao destacar que escolhas tecnológicas não são neutras, mas carregam implicações políticas que afetam a distribuição de poder na sociedade.[349]

Donald Maclennan e Judy Wajcman exploram a ideia de que as tecnologias são simultaneamente técnicas e sociais.[350] Eles enfatizam que o desenvolvimento tecnológico é moldado por relações de gênero, poder e

---

[344] PINCH, T. J.; BIJKER, W. E. The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other. Social Studies of Science, v. 14, n. 3, p. 399-441, 1984.
[345] BIJKER, W. E. Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change. Cambridge: MIT Press, 1995.
[346] BIJKER, W. E. Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change. Cambridge: MIT Press, 1995.
[347] WINNER, L. Do Artifacts Have Politics? Daedalus, v. 109, n. 1, p. 121-136, 1980
[348] WINNER, L. Do Artifacts Have Politics? Daedalus, v. 109, n. 1, p. 121-136, 1980
[349] WINNER, L. Do Artifacts Have Politics? Daedalus, v. 109, n. 1, p. 121-136, 1980
[350] MACLENNAN, D.; WAJCMAN, J. The Social Shaping of Technology. 2. ed. Buckingham: Open University Press, 1999.

cultura, e que a tecnologia, por sua vez, influencia essas relações.[351] Isso reforça a visão de que a tecnologia é uma construção social que reflete e afeta a organização social.

Aplicando a SCOT ao contexto da inteligência artificial (IA) e da discriminação algorítmica, percebe-se que os sistemas de IA são desenvolvidos dentro de contextos sociais específicos e refletem os valores, interesses e perspectivas dos grupos sociais envolvidos em sua criação.[352] Se as equipes de desenvolvimento são homogêneas em termos de gênero, raça ou classe social, é provável que os sistemas resultantes incorporem vieses que refletem essa homogeneidade.[353]

Os dados utilizados para treinar algoritmos de IA são coletados a partir de práticas sociais que podem conter preconceitos e desigualdades existentes.[354] A SCOT nos ajuda a entender que esses dados não são neutros, mas carregam significados sociais que influenciam o comportamento dos algoritmos. A definição dos problemas que a IA deve resolver e as soluções tecnológicas propostas estão intimamente ligadas às perspectivas dos grupos que controlam o desenvolvimento da tecnologia.[355]

Além disso, o conceito de flexibilidade interpretativa sugere que diferentes grupos podem interpretar e utilizar a IA de maneiras variadas. No entanto, quando o fechamento tecnológico ocorre sem a participação de todos

---

[351] MACLENNAN, D.; WAJCMAN, J. The Social Shaping of Technology. 2. ed. Buckingham: Open University Press, 1999.
[352] KITCHIN, R. Thinking Critically About and Researching Algorithms. Information, Communication & Society, v. 20, n. 1, p. 14-29, 2017.
[353] CRIADO PEREZ, C. Invisible Women: Data Bias in a World Designed for Men. New York: Abrams Press, 2019.
[354] NOBLE, S. U. Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press, 2018.
[355] WINNER, L. Do Artifacts Have Politics? Daedalus, v. 109, n. 1, p. 121-136, 1980

os grupos relevantes, tecnologias discriminatórias podem se tornar padrão.[356] Isso evidencia a necessidade de incluir diversas vozes no processo de desenvolvimento para evitar a marginalização de grupos vulneráveis.

Movimentos sociais, organizações de direitos civis e comunidades afetadas têm desempenhado um papel crucial ao questionar e desafiar as práticas discriminatórias na IA.[357] A SCOT enfatiza que os usuários não são apenas consumidores passivos, mas agentes ativos que podem influenciar a direção do desenvolvimento tecnológico. Essa interação dinâmica pode levar a revisões nos sistemas de IA para torná-los mais justos e inclusivos.

Portanto, a Construção Social da Tecnologia oferece uma estrutura teórica para compreender como a IA e a discriminação algorítmica são produtos de processos sociais complexos. Ao reconhecer que a tecnologia é moldada por escolhas sociais e políticas, torna-se possível identificar oportunidades para intervir e promover práticas de desenvolvimento que abordem vieses e promovam a equidade.[358]

9. Quebrando a Estrutura: IA como Ferramenta de Transformação

Embora a inteligência artificial tenha o potencial de reforçar estruturas de poder existentes, incluindo o patriarcalismo, ela também pode ser utilizada como uma ferramenta poderosa de transformação social. Para que isso ocorra, é fundamental desenvolver e treinar sistemas de IA com uma visão crítica de suas implicações éticas e sociais.[359] Essa abordagem permite que a IA não

[356] BIJKER, W. E. Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change. Cambridge: MIT Press, 1995.
[357] EUBANKS, V. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin's Press, 2018.
[358] OUDSHOORN, N.; PINCH, T. How Users Matter: The Co-Construction of Users and Technology. Cambridge: MIT Press, 2003.
[359] D'IGNACIO, C.; KLEIN, L. Data Feminism. Cambridge: MIT Press, 2020.

apenas evite reproduzir vieses históricos, mas ativamente contribua para a desconstrução de hierarquias opressivas.

9.1. A Importância da Plasticidade Cultural

Gilberto Freyre destaca a plasticidade cultural brasileira, referindo-se à capacidade de absorver, adaptar e transformar influências diversas para criar algo novo e único.[360] Essa característica pode ser relacionada ao potencial da IA de se moldar a diferentes contextos culturais e sociais quando desenvolvida com essa intenção. Ao promover a inclusão de diversas vozes no desenvolvimento da IA, estamos exercitando essa plasticidade cultural, permitindo que a tecnologia reflita a multiplicidade da sociedade brasileira.

A plasticidade cultural implica em uma flexibilidade e abertura para a incorporação de elementos heterogêneos, resultando em uma cultura dinâmica e resiliente. Aplicada à IA, essa ideia sugere que os sistemas inteligentes devem ser projetados para aprender e evoluir com base em uma ampla gama de dados culturais e sociais.[361] Isso não apenas melhora a eficiência dos algoritmos, mas também promove a criação de soluções mais inclusivas e representativas.

Além disso, essa flexibilidade cultural na IA pode facilitar a inovação, permitindo que tecnologias emergentes atendam a necessidades específicas de diferentes comunidades.[362] Ao reconhecer e valorizar a diversidade cultural, a IA pode contribuir para o desenvolvimento de aplicações que respeitem tradições locais e promovam o desenvolvimento sustentável.

---

[360] FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).
[361] SAMUEL, G. L. Cultural contexts in machine learning fairness. Ethics and Information Technology, v. 21, n. 1, p. 1-13, 2019.
[362] SAMBASIVAN, N.; HUTCHINSON, B.; RACE, A.; KALKER, T.; MUKHERJEE, S. Re-imagining algorithmic fairness in India and beyond. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. New York: ACM, 2021. p. 1-14.

Por fim, a plasticidade cultural na IA pode ajudar a mitigar os riscos de imperialismo cultural tecnológico, onde soluções padronizadas são impostas a contextos diversos.[363] Ao adaptar a tecnologia às especificidades culturais, evitamos a homogeneização e preservamos a riqueza das identidades locais, alinhando-se à visão de Freyre sobre a valorização da diversidade.

9.2. Convivência de Contrastes e Harmonia nas Diferenças

Freyre enfatiza a convivência de contrastes e a possibilidade de harmonia entre diferentes grupos étnicos e culturais. Essa visão pode ser aplicada ao desenvolvimento de IA que não apenas evita vieses, mas que também valoriza e promove a diversidade. Isso significa criar algoritmos que reconheçam e respeitem as particularidades de cada grupo social, contribuindo para uma sociedade mais inclusiva e equitativa.

A convivência de contrastes na IA envolve o desenvolvimento de sistemas que sejam capazes de lidar com dados heterogêneos e oferecer respostas que considerem múltiplas perspectivas.[364] Isso é especialmente importante em aplicações como processamento de linguagem natural, onde a consideração de dialetos e variações linguísticas pode melhorar significativamente a acessibilidade e a usabilidade.[365]

Promover harmonia nas diferenças também implica em garantir que a IA não privilegie um grupo em detrimento de outro. Técnicas de aprendizado de máquina justo (fair machine learning) podem ser implementadas para assegurar

[363] IRANI, L.; VERTESI, J.; DOURISH, P.; PHILIPS, K.; GRINTER, R. Postcolonial computing: A lens on design and development. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM, 2010. p. 1311-1320.
[364] MULLAINATHAN, S.; SPIES, D. Machine learning: An applied econometric approach. Journal of Economic Perspectives, v. 31, n. 2, p. 87-106, 2017.
[365] MULLAINATHAN, S.; SPIES, D. Machine learning: An applied econometric approach. Journal of Economic Perspectives, v. 31, n. 2, p. 87-106, 2017.

que os modelos não apresentem discriminações sistêmicas.[366] Dessa forma, a IA atua como um agente de equidade, refletindo a visão de Freyre sobre a possibilidade de coexistência harmoniosa entre diferentes grupos.

Além disso, a consideração das diferenças culturais e sociais no desenvolvimento de IA pode levar a soluções mais eficazes. Por exemplo, em sistemas de saúde, algoritmos que levam em conta variáveis socioeconômicas e culturais podem fornecer diagnósticos e tratamentos mais precisos para populações diversas.[367] Isso demonstra como a valorização da diversidade pode resultar em benefícios tangíveis para a sociedade.

9.3. A Família Patriarcal e a Modernização Tecnológica

A modernização tecnológica pode desafiar normas tradicionais, oferecendo oportunidades para a reconfiguração das dinâmicas familiares e sociais.[368] A IA pode contribuir para esse processo ao promover igualdade de oportunidades no mercado de trabalho, permitindo que todas as pessoas, independentemente de gênero, participem ativamente da vida profissional e familiar.

Empresas têm utilizado IA para eliminar vieses inconscientes nos processos de recrutamento. Ferramentas de IA estão sendo amplamente utilizadas para identificar disparidades salariais e promover a equidade dentro das organizações.[369] Grandes empresas, como Salesforce e Starbucks, estão adotando soluções baseadas em inteligência artificial para analisar dados

---

[366] ZHANG, B. H.; MITCHELL, M.; BAEZA-YATES, R. Fairness, accountability, and transparency in machine learning: Biases, methodologies, and challenges. AI Magazine, v. 39, n. 3, p. 55-66, 2018.
[367] OBERMEYER, Z.; POWELL, T.; ELEGBEDE, F.; MULLAINATHAN, S. Dissecting racial bias in an algorithm used to manage the health of populations. Science, v. 366, n. 6464, p. 447-453, 2019.
[368] CASTELLS, M. The Rise of the Network Society. 2. ed. Chichester: Wiley-Blackwell, 2010.
[369] CHURCH, Zach. Blendoor Battles Bias in Hiring. MIT Sloan, 2021. Disponível em: https://mitsloan.mit.edu/ideas-made-to-matter/blendoor-battles-bias-hiring . Acesso em: 10 out. 2023

salariais, identificar diferenças injustificadas e corrigir essas disparidades de maneira mais eficiente.[370] Essas ferramentas oferecem uma análise detalhada sobre a remuneração de funcionários com cargos e qualificações semelhantes, destacando desigualdades de gênero, raça e outros fatores que impactam a equidade no local de trabalho.[371]

Além de identificar as disparidades, essas tecnologias ajudam as organizações a entenderem as causas dessas desigualdades, promovendo ajustes significativos nas políticas de remuneração.[372] Através do uso dessas ferramentas, empresas não apenas corrigem as injustiças salariais, mas também criam uma cultura de transparência e igualdade, o que melhora a retenção de talentos e fortalece a reputação da organização no mercado.[373] O impacto positivo na cultura interna e nos resultados financeiros destaca o papel crucial que a IA pode desempenhar em iniciativas de justiça social, principalmente quando usada de forma estratégica e ética.[374]

Iniciativas como Women in AI e Black in AI trabalham para aumentar a representação de grupos sub-representados na área de inteligência artificial, promovendo diversidade e perspectivas variadas no desenvolvimento

---

[370] CHURCH, Zach. Blendoor Battles Bias in Hiring. MIT Sloan, 2021. Disponível em: https://mitsloan.mit.edu/ideas-made-to-matter/blendoor-battles-bias-hiring . Acesso em: 10 out. 2023

[371] CHURCH, Zach. Blendoor Battles Bias in Hiring. MIT Sloan, 2021. Disponível em: https://mitsloan.mit.edu/ideas-made-to-matter/blendoor-battles-bias-hiring . Acesso em: 10 out. 2023

[372] ENGELMEIER, Shirley. The App for Avoiding Gender Bias in Recruitment. Advisorpedia, 2021. Disponível em: https://www.advisorpedia.com/inclusion-initiative/blendoor-the-app-for-avoiding-gender-bias-in-recruitment/ . Acesso em: 10 out. 2023.

[373] ENGELMEIER, Shirley. The App for Avoiding Gender Bias in Recruitment. Advisorpedia, 2021. Disponível em: https://www.advisorpedia.com/inclusion-initiative/blendoor-the-app-for-avoiding-gender-bias-in-recruitment/ . Acesso em: 10 out. 2023.

[374] ENGELMEIER, Shirley. The App for Avoiding Gender Bias in Recruitment. Advisorpedia, 2021. Disponível em: https://www.advisorpedia.com/inclusion-initiative/blendoor-the-app-for-avoiding-gender-bias-in-recruitment/ . Acesso em: 10 out. 2023.

tecnológico.[375] Essas organizações oferecem programas de mentoria, networking e oportunidades de desenvolvimento profissional. A inclusão de mulheres e minorias é essencial para garantir que a IA reflita as necessidades e experiências de toda a sociedade. Isso evita a reprodução de vieses e promove a criação de soluções tecnológicas mais justas e equitativas, respondendo à necessidade apontada por Freyre de adaptação e evolução das estruturas sociais diante das mudanças tecnológicas.

Esses exemplos ilustram como a IA pode ser utilizada para desafiar estruturas patriarcais, promovendo igualdade de gênero tanto no ambiente profissional quanto no familiar. A modernização tecnológica, quando orientada por princípios de equidade, tem o potencial de reconfigurar relações sociais, conforme analisado por Freyre em suas obras. Freyre destacou a capacidade da sociedade brasileira de se adaptar e evoluir diante de mudanças econômicas e tecnológicas.[376] A IA representa uma oportunidade contemporânea para promover transformações sociais positivas, desde que desenvolvida e implementada com consciência crítica e inclusão.

9.4. Regionalismo e Valorização das Culturas Locais

Gilberto Freyre valorizou profundamente o regionalismo e a importância das culturas locais na formação da identidade nacional brasileira.[377] Ele argumentava que a riqueza cultural do país advém da diversidade de suas regiões, cada uma com suas peculiaridades em termos de linguagem, costumes e tradições. No contexto da inteligência artificial, essa perspectiva sugere a necessidade de desenvolver tecnologias que respeitem e promovam as

[375] WOMEN IN AI. About Us. 2021. Disponível em: https://www.womeninai.co/. Acesso em: 10 out. 2023; BLACK IN AI. About Black in AI. 2021. Disponível em: https://blackinai.github.io/ . Acesso em: 10 out. 2023.
[376] FREYRE, G. Sobrados e Mucambos. Rio de Janeiro: José Olympio, 1936.
[377] FREYRE, G. O Mundo que o Português Criou. Lisboa: Livros do Brasil, 1952.

particularidades regionais, evitando a homogeneização cultural imposta por soluções tecnológicas padronizadas.

A IA oferece ferramentas poderosas para preservar e valorizar as identidades locais. Algoritmos podem ser treinados para reconhecer variações linguísticas, expressões culturais e necessidades específicas de diferentes regiões, contribuindo para a inclusão digital e o fortalecimento das culturas regionais. Por exemplo, o projeto Vozes da Bahia utiliza técnicas de processamento de linguagem natural para reconhecer e interpretar o sotaque e as expressões idiomáticas típicas do estado da Bahia.[378] Ao adaptar assistentes virtuais e sistemas de reconhecimento de voz para compreender essas variações, o projeto facilita a interação de usuários locais com a tecnologia, respeitando suas particularidades linguísticas.[379]

Além disso, a IA tem sido utilizada para revitalizar línguas indígenas e tradicionais, auxiliando na educação e preservação cultural. O aplicativo Aikuma permite a gravação e tradução de histórias orais em línguas ameaçadas de extinção, utilizando técnicas de IA para facilitar a transcrição e o aprendizado.[380] No Brasil, iniciativas como a parceria entre o Google e a comunidade indígena Nheengatu resultaram na criação de ferramentas de tradução e preservação da

[378] SILVA, A. L.; SANTOS, B. R. Vozes da Bahia: Reconhecimento de fala com variação dialetal. Anais do Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana, p. 200-209, 2019.
[379] SILVA, A. L.; SANTOS, B. R. Vozes da Bahia: Reconhecimento de fala com variação dialetal. Anais do Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana, p. 200-209, 2019.
[380] BIRD, S; HANKE, F; ADAMS, O; LEE, H. Aikuma: A mobile app for collaborative language documentation. In: Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services. New York: ACM, 2015.

língua Nheengatu, contribuindo para a manutenção do patrimônio linguístico e cultural dessas comunidades.[381]

A promoção do regionalismo na IA também pode impulsionar o desenvolvimento econômico local. Tecnologias adaptadas às necessidades regionais estimulam a inovação e o empreendedorismo em diferentes partes do país, reduzindo desigualdades socioeconômicas.[382] Por exemplo, a startup brasileira Agrosmart utiliza IA para oferecer soluções de agricultura inteligente adaptadas às condições climáticas e de solo de diferentes regiões do Brasil. Ao fornecer ferramentas personalizadas para agricultores, a tecnologia contribui para o aumento da produtividade e sustentabilidade agrícola, impulsionando o desenvolvimento econômico regional.

Outra iniciativa relevante é o Porto Digital em Recife, um parque tecnológico que incentiva o desenvolvimento de tecnologias inovadoras com foco nas necessidades locais.[383] O ecossistema reúne empresas e startups que utilizam IA para solucionar desafios específicos da região Nordeste, promovendo o crescimento econômico e a geração de empregos.[384] Esse movimento alinha-se à valorização do regionalismo proposta por Freyre, ao estimular soluções tecnológicas que respeitam e potencializam as características culturais e econômicas locais.

---

[381] BAISI, Anna. A preservação de línguas indígenas através da tecnologia. 2023. Disponível em: http://ipol.org.br/a-preservacao-de-linguas-indigenas-atraves-da-tecnologia/ . Acesso em: 10 out. 2023.
[382] KITCHIN, R. The real-time city? Big data and smart urbanism. GeoJournal, v. 79, n. 1, p. 1-14, 2014.
[383] PORTO DIGITAL. Porto Digital: O Parque Tecnológico do Recife. 2021. Disponível em: https://www.portodigital.org/noticias/conheca-o-porto-digital-o-maior-parque-tecnologico-urbano-e-aberto-do-brasil . Acesso em: 10 out. 2023.
[384] PORTO DIGITAL. Porto Digital: O Parque Tecnológico do Recife. 2021. Disponível em: https://www.portodigital.org/noticias/conheca-o-porto-digital-o-maior-parque-tecnologico-urbano-e-aberto-do-brasil . Acesso em: 10 out. 2023.

A utilização da IA para reconhecer e respeitar as diversidades regionais reflete a plasticidade cultural brasileira destacada por Freyre. Ao incorporar elementos culturais variados nos sistemas tecnológicos, evita-se a imposição de padrões homogêneos que podem desconsiderar ou marginalizar identidades locais. Isso também promove uma relação mais próxima entre as comunidades e a tecnologia, aumentando a adoção e o engajamento com soluções inovadoras.

Além disso, a IA pode auxiliar na preservação de manifestações culturais regionais. Projetos que utilizam técnicas de visão computacional e aprendizado de máquina para digitalizar e catalogar artefatos, músicas, danças e outras expressões culturais locais contribuem para a conservação e difusão desse patrimônio.[385] Ao disponibilizar esse conteúdo em plataformas digitais, torna-se possível compartilhar a riqueza cultural regional com um público mais amplo, valorizando as contribuições de diferentes partes do país para a identidade nacional.

Em suma, a inteligência artificial, quando orientada para valorizar as culturas locais, pode atuar como uma ferramenta poderosa de inclusão e preservação cultural. Essa abordagem está em consonância com os conceitos de Gilberto Freyre sobre o regionalismo e a importância de reconhecer a diversidade como elemento fundamental na formação da identidade brasileira. Ao desenvolver tecnologias que respeitam e promovem as particularidades regionais, contribuímos para a construção de uma sociedade mais plural e equitativa, onde a inovação tecnológica serve aos interesses e necessidades de todos os segmentos da população.

---

[385] RODRIGUES, L. F.; FERREIRA, J. C. Preservação do patrimônio cultural imaterial através de tecnologias digitais. Revista Brasileira de Preservação Digital, v. 2, n. 1, p. 45-60, 2018.

9.5. Educação e Inteligência Artificial como Ferramentas de Transformação Social

Gilberto Freyre, em suas obras, enfatizou o papel fundamental da educação na transformação social, defendendo que o desenvolvimento de uma sociedade justa e equilibrada passa pela formação integral de seus indivíduos.[386] Ele acreditava que a educação deveria ser um instrumento para promover a inclusão, valorizar a diversidade cultural e combater as desigualdades sociais.[387] No contexto contemporâneo, a inteligência artificial surge como uma ferramenta poderosa que pode potencializar esse processo transformador, alinhando-se aos ideais freyreanos.

Freyre via a educação não apenas como transmissão de conhecimentos acadêmicos, mas como um meio de fomentar a consciência crítica e a responsabilidade social.[388] Ele defendia uma educação que refletisse a pluralidade cultural brasileira, incorporando as contribuições de diferentes grupos étnicos e sociais.[389] Essa perspectiva buscava romper com modelos eurocêntricos e elitistas, promovendo um sistema educacional inclusivo e democrático.

A integração da IA na educação pode ser vista como uma continuidade desse pensamento, na medida em que permite a criação de ambientes de aprendizagem mais equitativos e acessíveis. A tecnologia pode auxiliar na superação de barreiras geográficas e socioeconômicas, oferecendo recursos educacionais de qualidade a comunidades historicamente excluídas. Dessa

[386] FREYRE, G. Novo Mundo nos Trópicos. São Paulo: Editora Global, 1962.
[387] FREYRE, G. Novo Mundo nos Trópicos. São Paulo: Editora Global, 1962.
[388] FREYRE, G. Novo Mundo nos Trópicos. São Paulo: Editora Global, 1962.
[389] FREYRE, G. Novo Mundo nos Trópicos. São Paulo: Editora Global, 1962.

forma, a IA atua como um catalisador para a realização da visão de Freyre sobre a educação como motor de transformação social.

A IA pode ser utilizada para personalizar o aprendizado, adaptando-se ao ritmo e estilo de cada estudante, o que é particularmente benéfico para aqueles que enfrentam desafios no sistema educacional tradicional.[390] Plataformas educacionais como a Khan Academy utilizam algoritmos de IA para identificar lacunas no conhecimento dos alunos e fornecer recursos específicos para suprir essas necessidades.[391] Isso promove a inclusão e melhora os resultados educacionais, reduzindo disparidades e democratizando o acesso ao conhecimento.

A personalização do aprendizado reconhece a individualidade de cada estudante, alinhando-se à ideia de Freyre de valorização das diferenças culturais e sociais. Ao oferecer um ensino adaptado, a IA respeita os contextos específicos dos alunos, considerando fatores como linguagem, cultura e background socioeconômico. Isso contribui para a construção de uma educação mais humanizada e eficaz.

Além disso, a personalização pode estimular a motivação e o engajamento dos estudantes, uma vez que o conteúdo é apresentado de forma relevante e significativa para eles. Essa abordagem pode reduzir a evasão escolar e aumentar a autoestima dos alunos, promovendo um ciclo virtuoso de aprendizagem e desenvolvimento pessoal.

Além disso, a IA pode auxiliar na formação contínua de profissionais, oferecendo treinamento e desenvolvimento de habilidades em resposta às

[390] HOLMES, W.; et al. Ethics of AI in education: Towards a community-wide framework. International Journal of Artificial Intelligence in Education, v. 29, n. 4, p. 1-23, 2019.
[391] KHAN ACADEMY. Our Mission. 2021. Disponível em: https://www.khanacademy.org/about . Acesso em: 10 out. 2023.

demandas do mercado de trabalho em constante evolução.[392] Ferramentas como a Coursera e a Udacity utilizam algoritmos para recomendar cursos e programas personalizados, capacitando indivíduos a se adaptarem às mudanças tecnológicas e econômicas. A formação contínua é essencial em um mundo onde as tecnologias e práticas profissionais evoluem rapidamente. A IA pode facilitar o acesso a programas de capacitação, independentemente da localização geográfica ou condição socioeconômica dos indivíduos. Isso democratiza o desenvolvimento profissional e abre novas oportunidades de ascensão social.

A alfabetização digital e tecnológica é crucial para capacitar indivíduos a compreender e participar ativamente do desenvolvimento tecnológico.[393] Programas educacionais que ensinam programação e ciência de dados a grupos sub-representados, como o Programaê! no Brasil, promovem a inclusão digital e asseguram que os benefícios da IA sejam amplamente distribuídos. Isso reflete a valorização que Freyre dava à diversidade e à inclusão como pilares para uma sociedade mais justa.

A inclusão digital é um passo fundamental para evitar que a IA amplie as desigualdades existentes. Ao capacitar pessoas de diferentes origens para utilizar e desenvolver tecnologias, criamos um ecossistema mais diverso e representativo. Isso não apenas promove a justiça social, mas também enriquece o campo tecnológico com múltiplas perspectivas e experiências. Além disso, a alfabetização digital empodera os indivíduos, permitindo que eles sejam agentes ativos na sociedade da informação.

---

[392] NADEEM, M.; MAJEED, A.; MUSTAFA, M. T.; ZAHID, T.; ZULFIQAR, S. Integration of artificial intelligence in higher education: The future of teaching and learning. Pakistan Journal of Science, v. 72, n. 4, p. 227-233, 2020.
[393] TUFEKCI, Z. Twitter and Tear Gas: The Power and Fragility of Networked Protest. New Haven: Yale University Press, 2018.

A IA também pode ser um agente de transformação social, ao ser empregada para monitorar e avaliar políticas públicas, identificando áreas onde intervenções são necessárias para promover a equidade.[394] Por exemplo, algoritmos de análise de dados podem ser utilizados para mapear regiões com déficit em serviços básicos, como saúde e educação, permitindo que governos direcionem recursos de forma mais eficaz.

Além disso, a IA pode amplificar vozes sub-representadas, oferecendo plataformas para que narrativas diversas sejam ouvidas.[395] Ferramentas de processamento de linguagem natural podem ser utilizadas para traduzir e disseminar conteúdos em línguas indígenas ou dialetos regionais, promovendo a preservação cultural e a inclusão. O projeto Radio Garden, por exemplo, permite que pessoas de todo o mundo escutem rádios locais, difundindo culturas e perspectivas diversas. Essa iniciativa está em consonância com a valorização da miscigenação cultural e da diversidade presente na obra de Freyre.

Ao dar visibilidade a culturas e idiomas menos representados, a IA contribui para a preservação do patrimônio cultural e para a promoção do respeito à diversidade. Isso é particularmente importante em um mundo globalizado, onde há o risco de homogeneização cultural. A valorização das múltiplas identidades culturais enriquece o diálogo global e promove a compreensão mútua entre os povos.

Além disso, ao facilitar o acesso a conteúdos diversos, a IA pode combater estereótipos e preconceitos, promovendo a empatia e a solidariedade. Isso

---

[395] BENDER, E. M.; GEBRU, T.; MCMILLAN-MAJOR, A.; SHMITCHELL, S. On the dangers of stochastic parrots: Can language models be too big? In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. New York: ACM, 2021. p. 610-623.

reflete a visão de Freyre sobre a importância da convivência harmoniosa entre diferentes culturas e etnias na construção de uma sociedade plural e inclusiva.

A IA também pode apoiar movimentos sociais e organizações que trabalham pela justiça social, fornecendo ferramentas para análise de dados, mobilização e comunicação eficaz.[396] Plataformas como a *Ushahidi* utilizam IA para coletar e visualizar dados de crises humanitárias, conflitos e desastres naturais, permitindo uma resposta mais rápida e coordenada. A tecnologia pode potencializar o impacto de movimentos sociais, ajudando a organizar ações, disseminar informações e engajar a população. A IA facilita a identificação de tendências e necessidades emergentes, permitindo respostas mais ágeis e efetivas. Isso é crucial em contextos onde a rapidez de atuação pode salvar vidas ou prevenir injustiças.

Freyre acreditava na plasticidade cultural brasileira, na capacidade de absorver e transformar influências diversas em algo novo e único. A IA pode potencializar essa característica ao facilitar a troca de conhecimentos e experiências entre diferentes culturas, promovendo o intercâmbio e a inovação. Por meio de plataformas colaborativas e redes sociais inteligentes, é possível conectar indivíduos de origens distintas, enriquecendo o tecido social e cultural.

A tecnologia permite que ideias e práticas culturais sejam compartilhadas em escala global, promovendo a criatividade e a inovação resultantes da mistura de influências. Isso pode levar ao surgimento de soluções originais para problemas complexos, beneficiando a sociedade como um todo.

A crítica ao eurocentrismo feita por Freyre destaca a importância de reconhecer e valorizar as contribuições de diferentes povos na construção da

sociedade brasileira.[397] No desenvolvimento da IA, isso implica em evitar a reprodução de vieses culturais e assegurar que os sistemas tecnológicos reflitam a diversidade humana. Iniciativas como a Black in AI trabalham para aumentar a representatividade de pessoas negras na área de inteligência artificial, promovendo perspectivas variadas e combate ao preconceito.[398]

A diversidade nas equipes de desenvolvimento de IA é crucial para evitar a perpetuação de estereótipos e discriminações. Profissionais de diferentes origens podem identificar vieses que passariam despercebidos em grupos homogêneos, contribuindo para a criação de tecnologias mais justas e inclusivas.

Além disso, a valorização de perspectivas não ocidentais enriquece o campo da IA, introduzindo novas abordagens e soluções. Isso está alinhado com a visão de Gilberto Freyre sobre a importância de reconhecer a influência de múltiplas culturas na formação da identidade nacional e na construção de um futuro mais equitativo.

A interdisciplinaridade defendida por Freyre pode ser aplicada no desenvolvimento de IA, integrando conhecimentos das ciências humanas, sociais e exatas. Essa abordagem enriquece a compreensão dos impactos sociais e culturais da tecnologia, garantindo que ela seja desenvolvida de maneira ética e alinhada às necessidades humanas. Projetos como o AI for Social Good reúnem profissionais de diferentes áreas para criar soluções que abordem desafios sociais complexos.

---

[397] FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).
[398] BLACK IN AI. About Black in AI. 2021. Disponível em: https://blackinai.github.io/ . Acesso em: 10 out. 2023.

10. Estratégias Práticas para Melhorar Dados de Treinamento de IA e Reduzir Vieses Discriminatórios

A melhoria dos dados de treinamento de IA para torná-los menos enviesados é um processo multifacetado que requer esforços colaborativos entre desenvolvedores, pesquisadores, formuladores de políticas e comunidades. Ao implementar as estratégias mencionadas, é possível avançar em direção a sistemas de IA mais justos e equitativos, alinhados aos ideais de diversidade e inclusão. Esse compromisso coletivo é essencial para garantir que a IA contribua positivamente para a sociedade, promovendo igualdade e respeito a todos os indivíduos.

A inteligência artificial tornou-se parte integral de diversos aspectos da sociedade, influenciando desde decisões financeiras até recomendações em plataformas digitais. No entanto, a qualidade e a imparcialidade dos dados de treinamento são cruciais para garantir que os sistemas de IA não perpetuem vieses racistas, patriarcais, homofóbicos ou sexistas. Para construir sistemas mais justos e inclusivos, é essencial adotar estratégias práticas que melhorem esses dados.

10.1. Representatividade nos Dados de Treinamento

Uma das principais causas de vieses em sistemas de IA é a falta de representatividade nos dados de treinamento. Dados que refletem apenas um segmento da população podem levar a modelos que não generalizam bem para grupos diversos.[399] Isso ocorre porque os algoritmos aprendem padrões a partir

[399] BENDER, E. M.; GEBRU, T.; MCMILLAN-MAJOR, A.; SHMITCHELL, S. On the dangers of stochastic parrots: Can language models be too big? In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. New York: ACM, 2021. p. 610-623.

dos dados fornecidos; se esses dados são enviesados, o resultado será um sistema que reproduz e até amplifica tais vieses.[400]

Para mitigar esse problema, é fundamental coletar dados diversificados, ampliando as fontes para incluir diferentes gêneros, etnias, orientações sexuais e origens socioeconômicas.[401] Essa diversidade de dados permite que os modelos de IA reconheçam e atendam às necessidades de uma população mais ampla. Além disso, a inclusão de múltiplas perspectivas culturais e sociais enriquece a capacidade do sistema de compreender nuances e contextos variados.

É importante garantir equilíbrio nos conjuntos de dados, verificando se as categorias demográficas estão equilibradas e evitando que grupos minoritários sejam sub-representados.[402] A sub-representação pode levar a resultados prejudiciais, como o tratamento injusto de indivíduos pertencentes a esses grupos.[403] Técnicas estatísticas podem ser empregadas para analisar a distribuição dos dados e ajustar o conjunto de treinamento de forma a promover a equidade.[404]

Adicionalmente, estabelecer parcerias com organizações e comunidades que representam grupos sub-representados pode facilitar a coleta de dados mais inclusivos.[405] Essas parcerias podem fornecer insights valiosos sobre como

---

[400] BENDER, E. M.; GEBRU, T.; MCMILLAN-MAJOR, A.; SHMITCHELL, S. On the dangers of stochastic parrots: Can language models be too big? In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. New York: ACM, 2021. p. 610-623.
[401]BUOLAMWINI, J.; GEBRU, T. Gender shades: intersectional accuracy disparities in commercial gender classification. In: Conference on Fairness, Accountability and Transparency, 2018. Proceedings p.77-91.
[402] JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. Science, v. 349, n. 6245, p. 255-260, 2015.
[403] JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. Science, v. 349, n. 6245, p. 255-260, 2015.
[404] JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. Science, v. 349, n. 6245, p. 255-260, 2015.
[405] HOVY, D.; SPRUIT, S. L. The social impact of natural language processing. In: 54th Annual Meeting of the Association for Computational Linguistics, 2016. Proceedings... p.591-598.

melhor representar essas comunidades nos dados, garantindo que suas vozes e experiências sejam adequadamente incorporadas nos modelos de IA. Além disso, políticas de privacidade e consentimento ético devem ser rigorosamente aplicadas para proteger os direitos dos indivíduos durante a coleta de dados.[406]

Outra estratégia eficaz é o uso de técnicas de *data augmentation* específicas para ampliar a representatividade nos conjuntos de dados.[407] Essas técnicas permitem criar variações dos dados existentes, aumentando a diversidade sem a necessidade de coletar grandes volumes de novos dados. No entanto, é essencial que essas variações sejam geradas de maneira que respeite a autenticidade e evite introduzir novos vieses.

10.2. Processos de Anotação de Dados e Vieses Humanos

Os processos de anotação de dados também podem introduzir vieses se não forem conduzidos de forma cuidadosa. O treinamento de anotadores é essencial para capacitá-los a reconhecer e evitar vieses inconscientes, fornecendo diretrizes claras e exemplos de situações que podem levar a discriminações.[408] Os anotadores, sendo humanos, podem inadvertidamente projetar seus preconceitos nos dados, o que se refletirá no desempenho do sistema de IA.

Formar equipes diversas de anotadores traz múltiplas perspectivas ao processo, reduzindo a influência de preconceitos individuais.[409] A diversidade

---

[406] DIAZ, C.; TENE, O.; GÜRSES, S. Hero or villain: the data controller in privacy law and technologies. International Data Privacy Law, v.8, n.1, p.4-11, 2018.
[407] SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. Journal of Big Data, v.6, n.1, 2019.
[408] SNOW, Rion et al. Cheap and Fast—But is it Good? Evaluating Non-Expert Annotations for Natural Language Tasks. In: CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING (EMNLP), 2008, Honolulu. Proceedings... Stroudsburg: Association for Computational Linguistics, 2008. p. 254-263.
[409] FORT, K.; ADDA, G.; COHEN, K. B. Amazon Mechanical Turk: gold mine or coal mine? Computational Linguistics, v.37, n.2, p.413-420, 2011.

na equipe de anotação ajuda a equilibrar percepções e interpretações, minimizando o risco de vieses sistemáticos. Além disso, a rotatividade de anotadores e a revisão cruzada de anotações podem servir como mecanismos adicionais de controle de qualidade.

Implementar processos de dupla revisão, onde cada item é anotado por mais de uma pessoa, pode ajudar a identificar discrepâncias e vieses potenciais.[410] Quando há desacordo entre os anotadores, um processo de resolução pode ser ativado, envolvendo discussões ou até mesmo um terceiro anotador para desempatar. Isso assegura que as anotações sejam mais precisas e imparciais.

Além disso, a utilização de ferramentas de anotação assistida por IA pode auxiliar os anotadores, fornecendo sugestões e destacando possíveis áreas de vieses.[411] Essas ferramentas podem alertar os anotadores sobre decisões inconsistentes ou enviesadas, promovendo maior consistência e qualidade nas anotações. No entanto, é crucial que essas ferramentas sejam cuidadosamente projetadas para não introduzir novos vieses no processo.

10.3. Identificação e Correção de Vieses nos Dados

Antes de utilizar os dados para treinar modelos, é fundamental identificar e corrigir potenciais vieses. A análise estatística permite detectar desequilíbrios e padrões discriminatórios nos dados.[412] Por exemplo, técnicas de mineração de

[410] SHENG, V. S.; PROVOST, F.; IPEIROTIS, P. G. Get another label? Improving data quality and data mining using multiple, noisy labelers. In: 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2008. Proceedings... p.614-622.
[411] RATNER, A. et al. Snorkel: rapid training data creation with weak supervision. Proceedings of the VLDB Endowment, v.11, n.3, p.269-282, 2017.
[412] BAROCAS, S.; SELBST, A. D. Big data's disparate impact. California Law Review, v. 104, n. 3, p. 671-732, 2016.

dados podem ser aplicadas para identificar correlações indesejadas entre características sensíveis e variáveis de saída.

Ferramentas de visualização podem auxiliar na identificação de discrepâncias. Gráficos e dashboards interativos permitem aos analistas explorar os dados de forma intuitiva, revelando tendências que podem não ser evidentes em análises puramente numéricas. Isso facilita a detecção de áreas problemáticas que requerem intervenção.

Aplicar métodos de normalização e balanceamento, como reamostragem ou ponderação, ajuda a corrigir distribuições desiguais nos dados.[413] Essas técnicas asseguram que o modelo não seja indevidamente influenciado por categorias super-representadas. Além disso, a implementação de algoritmos de mitigação de vieses durante o pré-processamento dos dados pode reduzir ainda mais o risco de discriminação algorítmica.

Outra abordagem é a aplicação de técnicas de *fairness-aware data mining*, que visam desenvolver modelos que sejam justos em relação a atributos sensíveis.[414] Essas técnicas modificam os dados ou os algoritmos para minimizar a dependência entre a variável de decisão e os atributos sensíveis, sem comprometer significativamente a precisão do modelo.[415]

Adicionalmente, realizar testes piloto com os modelos em ambientes controlados pode ajudar a identificar possíveis impactos discriminatórios antes

[413] HE, Haibo; GARCIA, Edwardo A. Learning from Imbalanced Data. IEEE Transactions on Knowledge and Data Engineering, v. 21, n. 9, p. 1263-1284, 2009.
[414] KAMIRAN, F.; CALDERS, T. Data preprocessing techniques for classification without discrimination. Knowledge and Information Systems, v. 33, n. 1, p. 1-33, 2012.
[415] KAMIRAN, F.; CALDERS, T. Data preprocessing techniques for classification without discrimination. Knowledge and Information Systems, v. 33, n. 1, p. 1-33, 2012.

da implantação em larga escala.[416] Esses testes permitem ajustar o modelo e os dados de treinamento conforme necessário, assegurando que os vieses sejam corrigidos antecipadamente.

10.4. Desenvolvimento de Modelos com Consideração de Equidade

A incorporação de métricas de equidade permite avaliar a performance do modelo em diferentes grupos demográficos, garantindo que o desempenho seja consistente.[417] Métricas como igualdade de oportunidade e justiça estatística podem ser utilizadas para medir e comparar o desempenho entre grupos.

Utilizar algoritmos de aprendizado justo, projetados para minimizar vieses e incluir restrições de equidade durante o treinamento, também é uma prática recomendada.[418] Esses algoritmos ajustam o processo de aprendizado para considerar explicitamente a justiça como um objetivo, equilibrando a precisão preditiva com a necessidade de tratamento equitativo.

Além disso, a adoção de técnicas como o *adversarial debiasing* pode ser eficaz na redução de vieses nos modelos.[419] Essa abordagem utiliza redes adversárias para minimizar a capacidade do modelo de prever atributos sensíveis, incentivando-o a focar em características mais relevantes e não enviesadas.

É igualmente importante implementar auditorias regulares dos modelos após a implantação, monitorando continuamente o desempenho em relação às

---

[416] ENSIGN, D.; FRIEDLER, S. A.; NEVILLE, S.; SCHEIDEGGER, C.; VENKATASUBRAMANIAN, S. Runaway feedback loops in predictive policing. In: Proceedings of the 1st Conference on Fairness, Accountability and Transparency. PMLR, 2018. p. 160-171.
[417] VERMA, S.; RUBIN, J. Fairness definitions explained. In: 2018 IEEE/ACM International Workshop on Software Fairness, 2018. Proceedings... p.1-7.
[418] ZEMEL, R. et al. Learning fair representations. In: International Conference on Machine Learning, 2013. Proceedings... p.325-333.
[419] ZHANG, B. H.; LEMOINE, B.; MITCHELL, M. Mitigating unwanted biases with adversarial learning. In: AAAI/ACM Conference on AI, Ethics, and Society, 2018. Proceedings... p.335-340.

métricas de equidade.[420] Isso assegura que o modelo permaneça justo ao longo do tempo, especialmente à medida que novos dados são incorporados ou que o contexto operacional muda.

10.5. Transparência e Auditoria nos Processos de IA

A transparência nos processos de desenvolvimento e treinamento de IA é crucial para identificar e corrigir vieses. Manter documentação detalhada, com registros claros sobre a origem dos dados, métodos de coleta e processos de anotação, é essencial.[421] Essa prática não apenas facilita a detecção de problemas, mas também promove a confiança dos usuários e stakeholders nos sistemas de IA.

A adoção de "fichas de dados" para conjuntos de dados promove a transparência e facilita auditorias futuras. Essas fichas contêm informações detalhadas sobre as características do conjunto de dados, incluindo potenciais limitações e vieses conhecidos. Além disso, permitir que terceiros realizem auditorias externas nos sistemas de IA e seus dados de treinamento pode ajudar a identificar potenciais vieses não detectados internamente.[422]

A utilização de modelos interpretáveis ou explicáveis também contribui para a transparência.[423] Modelos de caixa branca permitem que desenvolvedores e usuários compreendam como as decisões são tomadas, facilitando a identificação de possíveis vieses. Ferramentas de *explainable AI*

[420] BANDY, J. Problematic machine behavior: a systematic literature review of algorithm audits. 2021. Disponível em: https://arxiv.org/abs/2102.04220. Acesso em: 10 out. 2023.
[421] GEBRU, T. et al. Datasheets for datasets. 2018. Disponível em: https://arxiv.org/abs/1803.09010 . Acesso em: 10 out. 2023.
[422] RAJI, I. D. et al. Saving face: investigating the ethical concerns of facial recognition auditing. In: AAAI/ACM Conference on AI, Ethics, and Society, 2020. Proceedings... p.145-151.
[423] RUDIN, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, v. 1, n. 5, p. 206-215, 2019.

(XAI) podem ser utilizadas para explicar as previsões dos modelos, aumentando a confiança e a responsabilidade.

Implementar políticas de governança de IA, que definam claramente os procedimentos para o desenvolvimento ético e transparente de sistemas de IA, é outra prática recomendada.[424] Essas políticas devem abranger aspectos como gestão de riscos, conformidade legal e ética, bem como mecanismos de responsabilização.

10.6. Educação e Sensibilização sobre Vieses em IA

Promover a consciência sobre os vieses em IA entre desenvolvedores e usuários é igualmente importante. Programas de treinamento que ofereçam cursos e workshops sobre ética em IA e vieses algorítmicos para profissionais da área são essenciais.[425] Esses programas capacitam os profissionais a identificar e mitigar vieses em suas práticas diárias.

Disponibilizar recursos educacionais que esclareçam como os vieses surgem e como podem ser mitigados contribui para uma maior sensibilização.[426] Materiais como guias, estudos de caso e ferramentas interativas podem auxiliar na disseminação desse conhecimento.

Incentivar a inclusão de disciplinas sobre ética e responsabilidade social nos currículos de ciência da computação e áreas afins pode preparar futuros profissionais para lidar com esses desafios.[427] A formação acadêmica que

[424] JOBIN, A.; IENCA, M.; VAYENA, E. The global landscape of AI ethics guidelines. Nature Machine Intelligence, v. 1, n. 9, p. 389-399, 2019.
[425] VILLARROEL, J. et al. Ethical considerations in artificial intelligence courses. In: IEEE Global Engineering Education Conference, 2019. Proceedings... p.892-897.
[426] WEST, S. M.; WHITTAKER, M.; CRAWFORD, K. Discriminating systems: Gender, race, and power in AI. AI Now Institute, 2019.
[427] FIESLER, C. Ethical considerations for research involving (speculative) public data. Proceedings of the ACM on Human-Computer Interaction, v.2, CSCW, p.1-13, 2018.

aborda os impactos sociais da tecnologia promove uma cultura de desenvolvimento responsável desde o início da carreira.

Além disso, promover discussões públicas e engajar a sociedade civil no debate sobre IA e vieses pode aumentar a conscientização geral.[428] Eventos, conferências e campanhas educativas podem envolver diferentes stakeholders, incluindo usuários finais, em uma reflexão coletiva sobre o uso ético da IA.

10.7. Políticas e Regulamentações para Práticas Responsáveis

Estabelecer diretrizes que incentivem práticas responsáveis é fundamental. A adoção de normas éticas que orientem o desenvolvimento de IA com foco em equidade e justiça proporciona um framework para decisões éticas.[429] Organizações podem implementar códigos de conduta e comitês de ética para supervisionar o cumprimento dessas diretrizes.

A implementação de regulamentação governamental que exija transparência algorítmica e responsabilização por danos causados por sistemas enviesados também é necessária.[430] Leis e políticas públicas podem estabelecer padrões mínimos para a proteção dos direitos dos indivíduos e promover a responsabilidade corporativa.

Incentivar a cooperação internacional na criação de padrões e regulamentos pode harmonizar as práticas em diferentes jurisdições, facilitando o desenvolvimento de IA ética em escala global.[431] Organizações internacionais podem desempenhar um papel chave na promoção de diretrizes universais.

---

[428]O'NEIL, C. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown, 2016.
[429] FLORIDI, L. et al. AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, v. 28, n. 4, p. 689-707, 2018.
[430] CATH, C. et al. Artificial intelligence and the 'good society': the US, EU, and UK approach. Science and Engineering Ethics, v.24, n.2, p.505-528, 2018.
[431] JOBIN, A.; IENCA, M.; VAYENA, E. The global landscape of AI ethics guidelines. Nature Machine Intelligence, v. 1, n. 9, p. 389-399, 2019.

Além disso, fomentar a pesquisa em ética e governança de IA através de financiamento e apoio institucional pode contribuir para o avanço de práticas responsáveis.[432] Investir em estudos que exploram os impactos sociais e éticos da IA ajuda a informar políticas eficazes e adaptadas aos desafios emergentes.

11. Conclusão

O estudo da interação entre as estruturas patriarcais e a inteligência artificial revela um campo fértil para análises críticas que combinam o pensamento de Gilberto Freyre com as questões contemporâneas. Ao explorar essa intersecção, torna-se evidente como as dinâmicas de poder e opressão descritas por Freyre persistem e se adaptam no contexto tecnológico atual. A IA não é uma entidade neutra; ela é moldada pelos valores, preconceitos e estruturas sociais existentes, o que a torna suscetível a reproduzir e amplificar desigualdades históricas.

Ao longo deste ensaio, foi possível traçar paralelos entre o patriarcado descrito por Freyre e as formas de exclusão e dominação que a IA pode perpetuar, especialmente no que diz respeito às desigualdades de gênero, raça e classe. Esses paralelos evidenciam que, assim como o patriarcado colonial influenciava todos os aspectos da vida social, a IA possui a capacidade de impactar diversas esferas da sociedade moderna. Os algoritmos, quando treinados em dados que refletem preconceitos sistêmicos, podem reforçar estereótipos e discriminações, afetando oportunidades de emprego, acesso a serviços e até mesmo decisões judiciais.

Os sistemas de IA, quando alimentados por dados que refletem uma sociedade marcada por essas hierarquias, tendem a reforçar esses padrões de

[432] MITTELSTADT, B. D. et al. The ethics of algorithms: mapping the debate. Big Data & Society, v.3, n.2, 2016.

opressão, naturalizando injustiças e reproduzindo desigualdades estruturais. Essa naturalização ocorre porque os algoritmos aprendem padrões existentes sem questioná-los, perpetuando vieses implícitos nos dados de treinamento. Além disso, a opacidade desses sistemas dificulta a identificação e correção de tais vieses, o que pode levar à legitimação de práticas discriminatórias sob a aparência de objetividade tecnológica.

A obra de Freyre, com sua análise detalhada sobre a formação social e cultural do Brasil, oferece um modelo teórico que nos ajuda a compreender como as tecnologias emergentes, como a IA, podem ser tanto ferramentas de emancipação quanto de opressão. Freyre enfatizava a complexidade das relações sociais e a influência das estruturas de poder nas interações humanas. Aplicando esse entendimento ao contexto da IA, reconhecemos que a tecnologia não é inerentemente boa ou má, mas reflete as intenções e valores de quem a desenvolve e utiliza.

Ao abordar as relações de poder que se perpetuam nas esferas digital e tecnológica, este ensaio destacou como a IA pode reproduzir o patriarcado, reforçando as estruturas de dominação e desigualdade que Freyre analisou no contexto colonial. Isso se manifesta na concentração de poder nas mãos de poucas corporações tecnológicas e na falta de diversidade nas equipes que desenvolvem essas ferramentas. A ausência de representatividade contribui para a criação de sistemas que não atendem às necessidades de todos os segmentos da população, excluindo vozes e experiências fundamentais para a construção de uma sociedade justa.

Entretanto, a IA também carrega em si o potencial de desconstruir essas mesmas hierarquias, desde que seja desenvolvida e implementada com um olhar crítico, ético e inclusivo. A tecnologia pode ser utilizada para identificar e

combater preconceitos, monitorar desigualdades e promover políticas públicas mais eficazes. Para isso, é essencial que haja um compromisso com a justiça social no processo de desenvolvimento da IA, envolvendo a participação de comunidades diversas e a consideração de contextos culturais variados.

A diversidade nas equipes de desenvolvimento, a transparência nos processos algorítmicos e a criação de políticas públicas que incentivem a equidade são medidas fundamentais para garantir que a tecnologia seja usada como uma ferramenta de transformação social. Essas ações contribuem para minimizar vieses, aumentar a confiabilidade dos sistemas e assegurar que os benefícios da IA sejam amplamente distribuídos. Além disso, a educação tecnológica inclusiva pode capacitar indivíduos de diferentes origens a participarem ativamente na criação e na crítica das tecnologias que os afetam.

Assim como Freyre sugeriu a importância de reconhecer e valorizar as múltiplas influências culturais na formação da sociedade brasileira, é essencial que o desenvolvimento da IA leve em consideração a pluralidade de vozes e experiências, evitando a imposição de padrões únicos que desconsiderem as diversidades. A incorporação de perspectivas variadas enriquece a tecnologia, tornando-a mais robusta e adaptável a diferentes realidades. Isso não apenas promove a inclusão, mas também fomenta a inovação, ao integrar conhecimentos e soluções provenientes de diversos contextos.

Portanto, ao integrar a reflexão freyriana com as críticas contemporâneas à IA, este ensaio buscou ressaltar a necessidade de uma abordagem interdisciplinar e ética na criação e utilização dessas tecnologias. Compreender a IA como um fenômeno social complexo exige a colaboração entre áreas como sociologia, antropologia, filosofia e ciência da computação. Essa integração

permite abordar os desafios éticos e sociais de forma abrangente, garantindo que a tecnologia sirva ao bem comum.

A inteligência artificial, longe de ser um fenômeno meramente técnico, é profundamente influenciada por questões sociais, culturais e políticas. Reconhecer essas influências e trabalhar para superá-las é essencial para que a IA possa ser uma força de inclusão, igualdade e justiça social. Isso implica em questionar as estruturas de poder existentes, promover a democratização do acesso à tecnologia e garantir a proteção dos direitos individuais e coletivos.

Contribuindo para a desconstrução de estruturas opressivas e a construção de uma sociedade mais equitativa e democrática, a IA tem o potencial de ser uma aliada na luta por justiça social. Para que isso se concretize, é imprescindível um compromisso coletivo com a ética, a transparência e a inclusão no desenvolvimento e aplicação dessas tecnologias. Inspirados pela obra de Gilberto Freyre, podemos vislumbrar uma trajetória onde a tecnologia atua como ferramenta de transformação positiva, refletindo a riqueza e a diversidade da experiência humana.

Referências

ADADI, A.; BERRADA, M. Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). IEEE Access, v. 6, p. 52138-52160, 2018.

AGGARWAL, C. C.; REDDY, C. K. Data Clustering: Algorithms and Applications. Boca Raton: CRC Press, 2013.

ALPAYDIN, E. Introduction to Machine Learning. 4. ed. Cambridge: MIT Press, 2020.

ANDREJEVIC, M. Automated Media. New York: Routledge, 2019.

ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica, 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing . Acesso em: 10 out. 2023.

AZEVEDO, C. O Patriarcado e a Construção da Sociedade Brasileira. São Paulo: Editora Contexto, 2004.

BAISI, Anna. A preservação de línguas indígenas através da tecnologia. 2023. Disponível em: http://ipol.org.br/a-preservacao-de-linguas-indigenas-atraves-da-tecnologia/ . Acesso em: 10 out. 2023.

BANDY, J. Problematic machine behavior: a systematic literature review of algorithm audits. 2021. Disponível em: https://arxiv.org/abs/2102.04220. Acesso em: 10 out. 2023.

BAROCAS, S.; SELBST, A. D. Big data's disparate impact. California Law Review, v. 104, n. 3, p. 671-732, 2016.

BARRETT, L. F. How Emotions Are Made: The Secret Life of the Brain. New York: Houghton Mifflin Harcourt, 2017.

BARTLETT, R.; MORRISON, S.; NELSON, L.; RISKER, M. Consumer Lending Discrimination in the FinTech Era. NBER Working Paper No. 25943, 2019.

BBC NEWS. Apple Card investigated after claims of gender bias. BBC News, 2019. Disponível em: https://www.bbc.com/news/business-50365609. Acesso em: 13 out. 2024.

BENDER, E. M.; GEBRU, T.; MCMILLAN-MAJOR, A.; SHMITCHELL, S. On the dangers of stochastic parrots: Can language models be too big? In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. New York: ACM, 2021. p. 610-623.

BENDICK, M.; EGAN, M. L. Changing workplace cultures to reduce employment discrimination. In: PLATT, J. (Ed.). Handbook of workplace diversity. Thousand Oaks: Sage Publications, 2000. p. 89-113.

BENJAMIN, R. Race After Technology: Abolitionist Tools for the New Jim Code. Cambridge: Polity Press, 2019.

BENNETT, W. L.; SEGERBERG, A. The logic of connective action: Digital media and the personalization of contentious politics. Information, Communication & Society, v. 15, n. 5, p. 739-768, 2012.

BERTRAND, M.; MULLAINATHAN, S. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. American Economic Review, v. 94, n. 4, p. 991-1013, 2004.

BHARGAVA, V.; LIVINGSTON, J. Regulating Algorithms in Consumer Credit. Harvard Journal of Law & Technology, v. 33, n. 1, p. 167-214, 2019.

BIJKER, W. E. Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change. Cambridge: MIT Press, 1995.

BIJKER, W. E.; HUGHES, T. P.; PINCH, T. J. The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology. Cambridge: MIT Press, 1987.

BIJKER, W. E.; LAW, J. Shaping Technology/Building Society: Studies in Sociotechnical Change. Cambridge: MIT Press, 1992.

BIRD, S; HANKE, F; ADAMS, O; LEE, H. Aikuma: A mobile app for collaborative language documentation. In: Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services. New York: ACM, 2015.

BLACK IN AI. About Black in AI. 2021. Disponível em: https://blackinai.github.io/ . Acesso em: 10 out. 2023.

BOLUKBASI, T.; CHANG, K.-W.; ZOU, J.; SALIGRAMA, V.; KALAI, A. Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In: Advances in Neural Information Processing Systems 29 (NIPS 2016). [S.l.: s.n.], 2016. p. 4349-4357.

BOSI, A. Dialética da Colonização. São Paulo: Companhia das Letras, 1992.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Diário Oficial da União, Brasília, DF, 15 ago. 2018.

BRYSON, Joanna; WINFIELD, Alan. Standardizing Ethical Design for Artificial Intelligence and Autonomous Systems. Computer, v. 50, n. 5, p. 116-119, 2017.

BUOLAMWINI, J.; GEBRU, T. Gender shades: intersectional accuracy disparities in commercial gender classification. In: Conference on Fairness, Accountability and Transparency, 2018. Proceedings p.77-91.

BURRELL, J. How the machine 'thinks': Understanding opacity in machine learning algorithms. Big Data & Society, v. 3, n. 1, p. 1-12, 2016.

BUTLER, J. Gender Trouble: Feminism and the Subversion of Identity. New York: Routledge, 1990.

CALDERS, T.; VERWER, S. Three naive Bayes approaches for discrimination-free classification. Data Mining and Knowledge Discovery, v.21, n.2, p.277-292, 2010.

CALISKAN, A.; BRYSON, J. J.; NARAYANAN, A. Semantics derived automatically from language corpora contain human-like biases. Science, v. 356, n. 6334, p. 183-186, 2017.

CÂMARA DOS DEPUTADOS. Projeto de Lei 21/2020. Dispõe sobre o uso da inteligência artificial no Brasil. Brasília, DF, 2020.

CASTELLS, M. A Sociedade em Rede. 2. ed. São Paulo: Paz e Terra, 2000.

CASTELLS, M. The Rise of the Network Society. 2. ed. Chichester: Wiley-Blackwell, 2010.

CASTELO, C. O Modo Português de Estar no Mundo: O Lusotropicalismo e a Ideologia Colonial Portuguesa (1933-1961). Porto: Afrontamento, 1998.

CATH, C. et al. Artificial intelligence and the 'good society': the US, EU, and UK approach. Science and Engineering Ethics, v.24, n.2, p.505-528, 2018.

CHURCH, Zach. Blendoor Battles Bias in Hiring. MIT Sloan, 2021. Disponível em: https://mitsloan.mit.edu/ideas-made-to-matter/blendoor-battles-bias-hiring. Acesso em: 10 out. 2023.

CITRON, D. K.; PASQUALE, F. The scored society: Due process for automated predictions. Washington Law Review, v. 89, n. 1, p. 1-33, 2014.

CORBETT-DAVIES, S.; GOEL, S. The measure and mismeasure of fairness: A critical review of fair machine learning. arXiv preprint arXiv:1808.00023, 2018.

COSTA, E. V. Da Monarquia à República: Momentos Decisivos. São Paulo: Editora Unesp, 1999.

COULDREY, N.; POWELL, A. Big Data from the South: The Beginning of a Conversation We Must Have. Television & New Media, v. 15, n. 5, p. 1-7, 2014.

COULDRY, N.; MEJIAS, U. A. The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism. Stanford: Stanford University Press, 2019.

CRAWFORD, K.; CALO, R. There is a blind spot in AI research. Nature, v. 538, n. 7625, p. 311-313, 2016.

CRENSHAW, K. Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. Stanford Law Review, v. 43, n. 6, p. 1241-1299, 1991.

CREVIER, D. AI: The Tumultuous History of the Search for Artificial Intelligence. New York: Basic Books, 1993.

CRIADO PEREZ, C. Invisible Women: Data Bias in a World Designed for Men. New York: Abrams Press, 2019.

CRUZ, F. A.; SILVA, R. M. Inteligência artificial e discriminação algorítmica: desafios para a regulação no Brasil. Revista de Direito, Estado e Telecomunicações, v. 12, n. 1, p. 1-22, 2020.

DAHLMAN, C.; MEALY, S.; WETZEL, D. China and the Digital Economy: Opportunities and Challenges. OECD Technical Report, 2018.

DASTIN, J. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters, 2018. Disponível em: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G . Acesso em: 10 out. 2023.

DAVIS, A. Mulheres, Raça e Classe. São Paulo: Boitempo, 2016.

DAY, M. Apple Card algorithm sparks gender bias allegations. Financial Times, 2019. Disponível em: https://www.ft.com/content/3d4b2d8e-05d3-11ea-a984-fbbacad9e7dd . Acesso em: 13 out. 2024.

DIAZ, C.; TENE, O.; GÜRSES, S. Hero or villain: the data controller in privacy law and technologies. International Data Privacy Law, v.8, n.1, p.4-11, 2018.

D'IGNACIO, C.; KLEIN, L. Data Feminism. Cambridge: MIT Press, 2020.

DOMINGOS, P. The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World. New York: Basic Books, 2015.

DOSHI-VELEZ, F.; KIM, B. Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608, 2017.

DWORK, C.; HARDT, M.; PITASSI, T.; REINGOLD, O.; ZEMEL, R. Fairness through awareness. In: Proceedings of the 3rd Innovations in Theoretical Computer Science Conference. New York: ACM, 2012. p. 214-226.

ENGELMEIER, Shirley. The App for Avoiding Gender Bias in Recruitment. Advisorpedia, 2021. Disponível em: https://www.advisorpedia.com/inclusion-initiative/blendoor-the-app-for-avoiding-gender-bias-in-recruitment/. Acesso em: 10 out. 2023.

ENSIGN, D.; FRIEDLER, S. A.; NEVILLE, S.; SCHEIDEGGER, C.; VENKATASUBRAMANIAN, S. Runaway feedback loops in predictive policing. In: Proceedings of the 1st Conference on Fairness, Accountability and Transparency. PMLR, 2018. p. 160-171.

EUBANKS, V. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin's Press, 2018.

EUROPEAN COMMISSION. White Paper on Artificial Intelligence: A European approach to excellence and trust. Bruxelas, 2020.

FAUSTO, B. História do Brasil. São Paulo: Edusp, 1995.

FEDERAL RESERVE. Report on the Economic Well-Being of U.S. Households. Board of Governors of the Federal Reserve System, 2017.

FENG, J.; JIN, C. The potential of cultural hybridity in the technology acceptance. International Journal of Innovation, Management and Technology, v. 1, n. 2, p. 194-198, 2010.

FERNANDES, F. A integração do negro na sociedade de classes. São Paulo: Ática, 1978.

FERNANDES, F. O Negro no Mundo dos Brancos. São Paulo: Difel, 1972.

FERRARA, E. et al. The Rise of Social Bots. Communications of the ACM, v. 59, n. 7, p. 96-104, 2016.

FIESLER, C. Ethical considerations for research involving (speculative) public data. Proceedings of the ACM on Human-Computer Interaction, v.2, CSCW, p.1-13, 2018.

FLORES, A. W.; BECHTEL, K.; LOWENKAMP, C. False Positives, False Negatives, and False Analyses. Federal Probation, v. 80, n. 2, p. 38-46, 2016.

FLORIDI, L. et al. AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, v. 28, n. 4, p. 689-707, 2018.

FORT, K.; ADDA, G.; COHEN, K. B. Amazon Mechanical Turk: gold mine or coal mine? Computational Linguistics, v.37, n.2, p.413-420, 2011.

FOUCAULT, M. A Arqueologia do Saber. Rio de Janeiro: Forense Universitária, 2012.

FOUCAULT, M. História da Sexualidade I: A Vontade de Saber. Rio de Janeiro: Graal, 1999.

FOUCAULT, M. Nascimento da Biopolítica: Curso no Collège de France (1978-1979). São Paulo: Martins Fontes, 2007.

FOUCAULT, M. Segurança, Território, População: Curso no Collège de France (1977-1978). São Paulo: Martins Fontes, 2008.

FOUCAULT, M. Vigiar e Punir: Nascimento da Prisão. Petrópolis: Vozes, 1975.

FREYRE, G. Casa-Grande & Senzala. 48. ed. São Paulo: Global Editora, 2006. (Original de 1933).

FREYRE, G. Lusotropicalismo. Lisboa: Livros do Brasil, 1952.

FREYRE, G. Novo Mundo nos Trópicos. São Paulo: Editora Global, 1962.

FREYRE, G. O Luso e o Trópico. Lisboa: Comissão Executiva das Comemorações do V Centenário da Morte do Infante D. Henrique, 1952.

FREYRE, G. O Mundo que o Português Criou. Lisboa: Livros do Brasil, 1952.

FREYRE, G. Ordem e Progresso. Rio de Janeiro: José Olympio, 1959.

FREYRE, G. Sobrados e Mucambos. Rio de Janeiro: José Olympio, 1936.

FREYRE, G. Tempo morto e outros tempos. Rio de Janeiro: José Olympio, 1981.

FRIEDLER, S. A. et al. A comparative study of fairness-enhancing interventions in machine learning. In: Conference on Fairness, Accountability, and Transparency, 2019. Proceedings... p.329-338.

FRIEDMAN, B.; NISSENBAUM, H. Bias in computer systems. ACM Transactions on Information Systems, v. 14, n. 3, p. 330-347, 1996.

GARVIE, C.; BEDOYA, A. M.; FRANKLE, J. The Perpetual Line-Up: Unregulated Police Face Recognition in America. Georgetown Law, Center on Privacy & Technology, 2016.

GEBRU, T. et al. Datasheets for datasets. 2018. Disponível em: https://arxiv.org/abs/1803.09010 . Acesso em: 10 out. 2023.

GILLESPIE, T. Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media. New Haven: Yale University Press, 2018.

GILPIN, L. H. et al. Explaining explanations: An overview of interpretability of machine learning. In: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA). IEEE, 2018. p. 80-89.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep Learning. Cambridge: MIT Press, 2016.

GOODMAN, B.; FLAXMAN, S. European Union regulations on algorithmic decision-making and a "right to explanation". AI Magazine, v. 38, n. 3, p. 50-57, 2017.

GOOGLE. Reducing gender bias in Google Translate. Google AI Blog, 2018. Disponível em: https://ai.googleblog.com/2018/12/reducing-gender-bias-in-google.html . Acesso em: 10 out. 2023.

GPAI. Global Partnership on Artificial Intelligence. 2020. Disponível em: https://gpai.ai/ . Acesso em: 10 out. 2023.

GRAY, M. L.; SURI, S. Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass. Boston: Houghton Mifflin Harcourt, 2019.

GUPTA, A.; ALLEN, L. A.; DHARMARAJAN, K.; RONG, J.; YEH, R. W.; KANNAN, S. G. Representation of women in trials of cardiovascular disease prevention. Circulation: Cardiovascular Quality and Outcomes, v. 11, n. 12, p. e004725, 2018.

HARARI, Y. N. 21 Lições para o Século 21. São Paulo: Companhia das Letras, 2018.

HARAWAY, D. J. Simians, Cyborgs, and Women: The Reinvention of Nature. New York: Routledge, 1991.

HARDY, S. et al. Mitigating Bias in AI-Enabled Face Recognition. IBM Journal of Research and Development, v. 63, n. 4/5, p. 3:1-3:9, 2019.

HASENBALG, C. Discriminação e Desigualdades Raciais no Brasil. Rio de Janeiro: Graal, 1979.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. The Elements of Statistical Learning. 2. ed. New York: Springer, 2009.

HE, Haibo; GARCIA, Edwardo A. Learning from Imbalanced Data. IEEE Transactions on Knowledge and Data Engineering, v. 21, n. 9, p. 1263-1284, 2009.

HEGEL, G. W. F. Fenomenologia do Espírito. Petrópolis: Vozes, 2012.

HICKS, M. Programmed Inequality: How Britain Discarded Women Technologists and Lost Its Edge in Computing. Cambridge: MIT Press, 2017.

HINCHLIFFE, E. Alexa, Siri and Gender Bias: Why Voice Assistants Need a Feminist Reboot. Forbes, 2019. Disponível em: https://www.forbes.com/sites/emilyhinchliffe/2019/05/22/alexa-siri-and-gender-bias-why-voice-assistants-need-a-feminist-reboot/ . Acesso em: 13 out. 2024.

HOLANDA, S. B. de. Raízes do Brasil. São Paulo: Companhia das Letras, 1995.

HOLMES, W.; et al. Ethics of AI in education: Towards a community-wide framework. International Journal of Artificial Intelligence in Education, v. 29, n. 4, p. 1-23, 2019.

HOLSTEIN, K.; WORTMAN VAUGHAN, J.; DUDIK, M.; WALLACH, H.; WORTMAN, S. Improving fairness in machine learning systems: What do industry practitioners need? In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. New York: ACM, 2019. p. 1-16.

HORKHEIMER, M.; ADORNO, T. W. Dialética do Esclarecimento. Rio de Janeiro: Zahar, 1985.

HOVY, D.; SPRUIT, S. L. The social impact of natural language processing. In: 54th Annual Meeting of the Association for Computational Linguistics, 2016. Proceedings... p.591-598.

HUNT, V.; PRINCE, S.; DIXON-FYLE, S.; YEE, L. Delivering through Diversity. McKinsey & Company, 2019.

HURLEY, M.; ADEBAYO, J. Credit scoring in the era of big data. Yale Journal of Law & Technology, v. 18, p. 148-216, 2016.

INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS (IEEE). Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems. New York: IEEE, 2019.

IRANI, L.; VERTESI, J.; DOURISH, P.; PHILIPS, K.; GRINTER, R. Postcolonial computing: A lens on design and development. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM, 2010. p. 1311-1320.

JAMES, L. DFS to Investigate Apple Card for Gender Discrimination. New York State Department of Financial Services, 2019. Disponível em: https://www.dfs.ny.gov/reports_and_publications/press_releases/pr1911101 . Acesso em: 13 out. 2024.

JASANOFF, S. States of Knowledge: The Co-Production of Science and Social Order. London: Routledge, 2004.

JOBIN, A.; IENCA, M.; VAYENA, E. The global landscape of AI ethics guidelines. Nature Machine Intelligence, v. 1, n. 9, p. 389-399, 2019.

JOHNSON, D. G. Technology with no human responsibility? Journal of Business Ethics, v. 127, n. 4, p. 707-715, 2015.

JOHNSON, D. G.; WETMORE, J. M. Technology and Society: Building Our Sociotechnical Future. Cambridge: MIT Press, 2009.

JOHNSON, S. Using AI to Organize Your Family Life. Family Tech Weekly, 2019.

JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. Science, v. 349, n. 6245, p. 255-260, 2015.

KAMIRAN, F.; CALDERS, T. Data preprocessing techniques for classification without discrimination. Knowledge and Information Systems, v. 33, n. 1, p. 1-33, 2012.

KELLEHER, J. D.; NAMEE, B. M.; DOOLEY, A. Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies. Cambridge: MIT Press, 2015.

KELTY, Christopher M. Two Bits: The Cultural Significance of Free Software. Durham: Duke University Press, 2008.

KHAN ACADEMY. Our Mission. 2021. Disponível em: https://www.khanacademy.org/about. Acesso em: 10 out. 2023.

KIRCHNER, L. When Discrimination Is Baked Into Algorithms. The Atlantic, 2016. Disponível em: https://www.theatlantic.com/business/archive/2016/09/discrimination-algorithms-disparate-impact/501329/ . Acesso em: 13 out. 2024.

KITCHIN, R. The real-time city? Big data and smart urbanism. GeoJournal, v. 79, n. 1, p. 1-14, 2014.

KITCHIN, R. Thinking Critically About and Researching Algorithms. Information, Communication & Society, v. 20, n. 1, p. 14-29, 2017.

KLEIN, H. K.; KLEINMAN, D. L. The Social Construction of Technology: Structural Considerations. Science, Technology, & Human Values, v. 27, n. 1, p. 28-52, 2002.

KLEINBERG, J.; MULLAINATHAN, S.; RAGHAVAN, M. Inherent Trade-Offs in the Fair Determination of Risk Scores. In: Proceedings of Innovations in Theoretical Computer Science Conference (ITCS), 2018.

KLINE, R.; PINCH, T. Users as Agents of Technological Change: The Social Construction of the Automobile in the Rural United States. Technology and Culture, v. 37, n. 4, p. 763-795, 1996.

KOSTKA, G. China's Social Credit Systems and Public Opinion: Explaining High Levels of Approval. New Media & Society, v. 21, n. 7, p. 1565-1593, 2019.

KUHN, T. S. A Estrutura das Revoluções Científicas. São Paulo: Perspectiva, 1962.

LATOUR, B. Reassembling the Social: An Introduction to Actor-Network-Theory. Oxford: Oxford University Press, 2005.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. Nature, v. 521, n. 7553, p. 436-444, 2015.

LEHMANN, J. F. et al. Facial recognition technologies in the wild: A call for a federal office. Science, v. 365, n. 6455, p. 128-130, 2019.

LESEMAN, S.; SEIDMAN, D. Facial Recognition Technology: Potential Risks and Policy Considerations. Congressional Research Service, 2019.

LEWIS, J. E. et al. Indigenous Protocol and Artificial Intelligence Position Paper. Honolulu, HI: The Initiative for Indigenous Futures and the Canadian Institute for Advanced Research, 2020.

LEWIS, P.; HENDERSEN, M.; BURKE, J. et al. Revealed: Facebook's internal rulebook on sex, terrorism and violence. The Guardian, 2018. Disponível em: https://www.theguardian.com/news/2018/may/21/revealed-facebook-internal-rulebook-sex-terrorism-violence. Acesso em: 10 out. 2023.

LIANG, F. et al. Constructing a Data-Driven Society: China's Social Credit System as a State Surveillance Infrastructure. Policy & Internet, v. 10, n. 4, p. 415-453, 2018.

LOURES, L.; SILVA, D. Regionalisms in natural language processing: A systematic literature review. Journal of Language and Literature, v. 11, n. 2, p. 27-34, 2020.

LUKÁCS, G. História e Consciência de Classe. São Paulo: Martins Fontes, 2012.

LUNDBERG, S. M.; LEE, S.-I. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems. 2017. p. 4765-4774.

MACLENNAN, D.; WAJCMAN, J. The Social Shaping of Technology. 2. ed. Buckingham: Open University Press, 1999.

MAGER, M.; PLATA, M.; LOPEZ, L. G.; MEZA, I. S.; HERNANDEZ, R. I. Challenges of language technologies for the indigenous languages of the Americas. In: Proceedings of the 27th International Conference on Computational Linguistics. Stroudsburg: ACL, 2018. p. 55-69.

MARCUSE, H. One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society. Boston: Beacon Press, 1964.

MARX, K. Manuscritos Econômico-Filosóficos. São Paulo: Boitempo, 2010.

MARX, K. O Capital: Crítica da Economia Política. São Paulo: Nova Cultural, 1988.

MARX, K.; ENGELS, F. A Ideologia Alemã. São Paulo: Boitempo, 2007.

MARX, Karl. Manuscritos Econômico-Filosóficos. São Paulo: Boitempo, 2010.

MARX, Karl. O Capital: Crítica da Economia Política. São Paulo: Nova Cultural, 1988.

MARX, Karl; ENGELS, Friedrich. A Ideologia Alemã. São Paulo: Boitempo, 2007.

MATTOS, H. C. B. Escravidão e Cidadania no Brasil Monárquico. Rio de Janeiro: Zahar, 2013.

MCCARTHY, J. et al. A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. AI Magazine, v. 27, n. 4, p. 12-14, 2006.

MEDEIROS, M.; ALBERTI, V. R. (Org.). Desigualdades Sociais, Violências e Racismo no Brasil. Rio de Janeiro: Editora FGV, 2020.

MITCHELL, M. et al. Model cards for model reporting. In: Conference on Fairness, Accountability, and Transparency, 2019. Proceedings... p.220-229.

MITCHELL, Margaret et al. Model Cards for Model Reporting. In: CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (FAT), 2019, Atlanta. Proceedings... New York: ACM, 2019. p. 220-229.

MITCHELL, T. Machine Learning. New York: McGraw-Hill, 1997.

MITTELSTADT, B. D. et al. The ethics of algorithms: mapping the debate. Big Data & Society, v.3, n.2, 2016.

MOHANTY, C. T. Under western eyes: Feminist scholarship and colonial discourses. Feminist Review, n. 30, p. 61-88, 1988.

MULLAINATHAN, S.; SPIES, D. Machine learning: An applied econometric approach. Journal of Economic Perspectives, v. 31, n. 2, p. 87-106, 2017.

MUNANGA, K. Rediscutindo a Mestiçagem no Brasil: Identidade Nacional versus Identidade Negra. Petrópolis: Vozes, 1999.

MURPHY, K. P. Machine Learning: A Probabilistic Perspective. Cambridge: MIT Press, 2012.

NADEEM, M.; MAJEED, A.; MUSTAFA, M. T.; ZAHID, T.; ZULFIQAR, S. Integration of artificial intelligence in higher education: The future of teaching and learning. Pakistan Journal of Science, v. 72, n. 4, p. 227-233, 2020.

NEFF, G.; NAGY, P. Talking to bots: Symbiotic agency and the case of Tay. International Journal of Communication, v. 10, p. 4915-4931, 2016.

NICKERSON, R. S. Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology, v. 2, n. 2, p. 175-220, 1998.

NIKOLIĆ, I.; HOLM, E. Agent-Based Modelling of Socio-Technical Systems. Dordrecht: Springer, 2019.

NOBLE, S. U. Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press, 2018.

NOGUEIRA, O. Preconceito de marca: as relações raciais em Itapetininga. São Paulo: Edusp, 2003.

NORMAN, D. A. The Design of Everyday Things. Revised and expanded edition. New York: Basic Books, 2013.

OBERMEYER, Z.; POWELL, T.; ELEGBEDE, F.; MULLAINATHAN, S. Dissecting racial bias in an algorithm used to manage the health of populations. Science, v. 366, n. 6464, p. 447-453, 2019.

OLTEANU, A. et al. Social data: Biases, methodological pitfalls, and ethical boundaries. Frontiers in Big Data, v. 2, p. 13, 2019.

OLTEANU, A.; CASTILLO, C.; DIAKOPOULOS, N.; ECHENIQUE, M. Social data: Biases, methodological pitfalls, and ethical boundaries. Frontiers in Big Data, v. 2, p. 13, 2019.

O'NEIL, C. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown, 2016.

ORTIZ, R. Cultura brasileira e identidade nacional. São Paulo: Brasiliense, 2002.

OUDSHOORN, N.; PINCH, T. How Users Matter: The Co-Construction of Users and Technology. Cambridge: MIT Press, 2003.

PAGE, S. E. The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies. Princeton: Princeton University Press, 2007.

PARISER, E. The Filter Bubble: What the Internet is Hiding from You. New York: Penguin Press, 2011.

PARRY, D. et al. Understanding the Limitations of AI: When Tay Tweets Go Bad. In: ACM CHI Conference on Human Factors in Computing Systems, 2016.

PASQUALE, F. The Black Box Society: The Secret Algorithms That Control Money and Information. Cambridge: Harvard University Press, 2015.

PINCH, T. J.; BIJKER, W. E. The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other. Social Studies of Science, v. 14, n. 3, p. 399-441, 1984.

PORTO DIGITAL. Porto Digital: O Parque Tecnológico do Recife. 2021. Disponível em: https://www.portodigital.org/noticias/conheca-o-porto-digital-o-maior-parque-tecnologico-urbano-e-aberto-do-brasil . Acesso em: 10 out. 2023.

PRADO, M. Formação do Brasil Contemporâneo. São Paulo: Brasiliense, 1994.

PRATES, M. O. R.; AVILA, S.; LAMB, L. C. Assessing gender bias in machine translation: A case study with Google Translate. Neural Computing and Applications, v. 32, n. 10, p. 6363-6381, 2020.

PRATES, M. O. R.; AVILA, S.; LAMB, L. C. Assessing Gender Bias in Machine Translation: A Case Study with Google Translate. arXiv preprint arXiv:1809.02208, 2018.

PROVOST, F.; FAWCETT, T. Data science and its relationship to big data and data-driven decision making. Big Data, v. 1, n. 1, p. 51-59, 2013.

RAJI, I. D. et al. Saving face: investigating the ethical concerns of facial recognition auditing. In: AAAI/ACM Conference on AI, Ethics, and Society, 2020. Proceedings... p.145-151.

RAJI, Inioluwa Deborah et al. Saving Face: Investigating the Ethical Concerns of Facial Recognition Auditing. In: AAAI/ACM CONFERENCE ON AI, ETHICS, AND SOCIETY (AIES), 2020, Nova Iorque. Proceedings... New York: ACM, 2020. p. 145-151.

RATNER, A. et al. Snorkel: rapid training data creation with weak supervision. Proceedings of the VLDB Endowment, v.11, n.3, p.269-282, 2017.

RAUHOFER, J.; VORNEFELD, B. Algorithmic discrimination and the European Union's General Data Protection Regulation. In: HIJMANS, H.; SCIROCCO, A. (Ed.). Data Protection and Privacy: The Internet of Bodies. Dordrecht: Springer, 2019. p. 51-74.

RIBEIRO, M. T.; SINGH, S.; GUESTRIN, C. "Why should I trust you?": Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016. p. 1135-1144.

ROBERTS, S. T. Behind the Screen: Content Moderation in the Shadows of Social Media. New Haven: Yale University Press, 2019.

RODRIGUES, L. F.; FERREIRA, J. C. Preservação do patrimônio cultural imaterial através de tecnologias digitais. Revista Brasileira de Preservação Digital, v. 2, n. 1, p. 45-60, 2018.

RUDIN, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, v. 1, n. 5, p. 206-215, 2019.

RUSSELL, S.; NORVIG, P. Artificial Intelligence: A Modern Approach. 3. ed. Upper Saddle River: Prentice Hall, 2016.

SAMBASIVAN, N.; HUTCHINSON, B.; RACE, A.; KALKER, T.; MUKHERJEE, S. Re-imagining algorithmic fairness in India and beyond. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. New York: ACM, 2021. p. 1-14.

SAMUEL, G. L. Cultural contexts in machine learning fairness. Ethics and Information Technology, v. 21, n. 1, p. 1-13, 2019.

SANTOS, B. S. Um Discurso sobre as Ciências. São Paulo: Cortez, 2008.

SANTOS, M. A.; PEREIRA, S. R.; OLIVEIRA, G. M. Aplicação de inteligência artificial na saúde pública regional. Revista de Saúde Pública, v. 54, p. 1-10, 2020.

SCHMIDHUBER, J. Deep learning in neural networks: An overview. Neural Networks, v. 61, p. 85-117, 2015.

SCHNEIDER, N. Decentralization: An Urgent Opportunity to Design Alternative Platforms for the Internet. Journal of Design and Science, 2018.

SCHUMAHER, S.; BRAZIL, E. Dicionário Mulheres do Brasil: De 1500 Até a Atualidade. Rio de Janeiro: Zahar, 2000.

SELBST, A. D. et al. Fairness and abstraction in sociotechnical systems. In: Proceedings of the Conference on Fairness, Accountability, and Transparency. 2019. p. 59-68.

SEVCENKO, N. Literatura como Missão: Tensões Sociais e Criação Cultural na Primeira República. São Paulo: Companhia das Letras, 1998.

SHAPIN, S.; SCHAFFER, S. Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life. Princeton: Princeton University Press, 1985.

SHENG, V. S.; PROVOST, F.; IPEIROTIS, P. G. Get another label? Improving data quality and data mining using multiple, noisy labelers. In: 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2008. Proceedings... p.614-622.

SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. Journal of Big Data, v.6, n.1, 2019.

SIDDHARTHA, V. Apple Card's Gender Bias: What Really Happened. Forbes, 2019. Disponível em: https://www.forbes.com/sites/siddharthavarma/2019/11/11/apple-cards-gender-bias-what-really-happened/ . Acesso em: 13 out. 2024.

SILVA, A. L.; SANTOS, B. R. Vozes da Bahia: Reconhecimento de fala com variação dialetal. Anais do Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana, p. 200-209, 2019.

SILVA, M. L.; FIGUEIREDO, D. R. Diversidade na área de computação no Brasil: uma análise de gênero e raça. Revista Brasileira de Informática na Educação, v. 29, p. 529-551, 2021.

SILVA, T. Inteligência Artificial e Sociedade: Reflexões Interdisciplinares. São Paulo: Editora da Universidade, 2020.

SISMONDO, S. An Introduction to Science and Technology Studies. Chichester: Wiley-Blackwell, 2010.

SLOAN, Luke; WRIGHT, Samuel (Ed.). Researching Social Media. London: SAGE Publications, 2015.

SMITH, A. The impact of technology on work-life balance. Journal of Business and Economics, v. 12, n. 3, p. 45-56, 2020.

SNOW, Rion et al. Cheap and Fast—But is it Good? Evaluating Non-Expert Annotations for Natural Language Tasks. In: CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING (EMNLP), 2008, Honolulu. Proceedings... Stroudsburg: Association for Computational Linguistics, 2008. p. 254-263.

SOUZA, L. M. A. *O Diálogo entre Gilberto Freyre e Franz Boas: Antropologia, História e Política*. Recife: Editora Massangana, 2002.

STRADER, M. K.; KABAT-ZINN, J. Mindfulness-based stress reduction for health care professionals: Results from a randomized trial. International Journal of Stress Management, v. 25, n. 2, p. 107-120, 2018.

SUTTON, R. S.; BARTO, A. G. Reinforcement Learning: An Introduction. 2. ed. Cambridge: MIT Press, 2018.

TATMAN, R. Gender and dialect bias in YouTube's automatic captions. In: Proceedings of the First ACL Workshop on Ethics in Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2017. p. 53-59.

TUFEKCI, Z. Twitter and Tear Gas: The Power and Fragility of Networked Protest. New Haven: Yale University Press, 2018.

TURNER, A. L. et al. Student Loan Default Rates among African American Students at For-Profit Colleges and Universities. Journal of Negro Education, v. 87, n. 1, p. 2-17, 2018.

UN REPORT. Report of the Panel of Experts on the Illegal Exploitation of Natural Resources and Other Forms of Wealth of the Democratic Republic of the Congo. United Nations, 2001.

UNESCO. I'd blush if I could: closing gender divides in digital skills through education. Paris: UNESCO, 2019.

UNESCO. Recommendation on the Ethics of Artificial Intelligence. Paris: UNESCO, 2021.

UNIÃO EUROPEIA. Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Bruxelas, 2021.

UNIÃO EUROPEIA. White Paper on Artificial Intelligence: A European Approach to Excellence and Trust. Bruxelas, 2020.

VAIDHYANATHAN, S. Antisocial Media: How Facebook Disconnects Us and Undermines Democracy. New York: Oxford University Press, 2018.

VEALE, M.; BRASS, I. Critical approaches to AI policy and governance. Internet Policy Review, v. 8, n. 2, p. 1-17, 2019.

VERMA, S.; RUBIN, J. Fairness definitions explained. In: 2018 IEEE/ACM International Workshop on Software Fairness, 2018. Proceedings... p.1-7.

VIGDOR, Neil. Apple Card Investigated After Gender Discrimination Complaints. The New York Times, 10 nov. 2019. Disponível em: https://www.nytimes.com/2019/11/10/business/Apple-credit-card-investigation.html . Acesso em: 13 out. 2024.

VILLARROEL, J. et al. Ethical considerations in artificial intelligence courses. In: IEEE Global Engineering Education Conference, 2019. Proceedings... p.892-897.

VINCENT, J. Microsoft's racist chatbot returns with new rules. The Verge, 2017. Disponível em: https://www.theverge.com/2017/8/8/16113530/microsoft-zo-ai-chatbot-returns . Acesso em: 13 out. 2024.

VOIGT, P.; BUSSCHE, A. The EU General Data Protection Regulation (GDPR): A Practical Guide. Springer International Publishing, 2017.

WA THIONG'O, N. Decolonising the Mind: The Politics of Language in African Literature. London: James Currey, 1986.

WAJCMAN, J. (2004). Technofeminism. Cambridge: Polity Press.

WELLER, A. Transparency: Motivations and challenges. In: ZHANG, Y.; ADALI, S. (Ed.). Transparent Data Mining for Big and Small Data. Cham: Springer, 2019. p. 23-40.

WEST, S. M.; KRAUT, R.; CHEW, H. E. I'd blush if I could: Closing gender divides in digital skills through education. UNESCO, 2019. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000367416 . Acesso em: 10 out. 2023.

WEST, S. M.; WHITTAKER, M.; CRAWFORD, K. Discriminating systems: Gender, race, and power in AI. AI Now Institute, 2019.

WHITTAKER, M. et al. AI now report 2018. New York: AI Now Institute at New York University, 2018.

WINNER, L. Do Artifacts Have Politics? Daedalus, v. 109, n. 1, p. 121-136, 1980.

WOMEN IN AI. About Us. 2021. Disponível em: https://www.womeninai.co/. Acesso em: 10 out. 2023.

ZEMEL, R. et al. Learning fair representations. In: International Conference on Machine Learning, 2013. Proceedings... p.325-333.

ZHANG, B. H.; LEMOINE, B.; MITCHELL, M. Mitigating unwanted biases with adversarial learning. In: AAAI/ACM Conference on AI, Ethics, and Society, 2018. Proceedings... p.335-340.

ZHANG, B. H.; MITCHELL, M.; BAEZA-YATES, R. Fairness, accountability, and transparency in machine learning: Biases, methodologies, and challenges. AI Magazine, v. 39, n. 3, p. 55-66, 2018.

ZLIOBAITĖ, I. Measuring discrimination in algorithmic decision making. Data Mining and Knowledge Discovery, v. 31, n. 4, p. 1060-1089, 2017.

ZUBOFF, S. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs, 2019.

ZUCKERMAN, E. Q&A: Ethan Zuckerman on the Failure of Facebook's 'Good Intentions'. MIT News, 2019. Disponível em: http://news.mit.edu/2019/qa-ethan-zuckerman-facebook-good-intentions-0312 . Acesso em: 13 out. 2024.